Alma Mater Studiorum
Università degli studi di Bologna

DIPARTIMENTO DI FISICA E ASTRONOMIA
Dottorato di Ricerca in Astronomia
CICLO XXVII

X-ray properties and evolution of high-redshift AGN, and the gas content of host galaxies

Dottorando
Fabio Vito

Relatori
Prof. Cristian Vignali

Dott. Roberto Gilli
Prof. Andrea Comastri

Coordinatore
Prof. L. Moscardini

Esame finale anno 2014

Settore Concorsuale: 02/C1 – Astronomia, Astrofisica, Fisica della Terra e dei Pianeti
Settore Scientifico-Disciplinare: FIS/05 – Astronomia e Astrofisica

"Astronomy compels the soul to look upward,
and leads us from this world to another."
Plato

"Allons-y!"
10^{th} *Doctor*

Contents

Executive summary

The presence of Super Massive Black Holes (SMBHs) with masses exceeding 10^9 $M_\odot$ at $z \gtrsim 7$ (e.g. Mortlock et al. 2011), when the Universe was less than 1 Gyr old, challenges our knowledge of SMBH formation and growth. To make a billion solar mass black hole in a few 10^8 years, continuous Eddington-limited, or even super-Eddington, accretion is required by the two most promising models of SMBH seeds formation, i.e. remnant of POP III stars (e.g. Volonteri & Begelman 2010) and direct collapse of massive gas clouds (e.g. Begelman 2010). Assessing the cosmic evolution of the AGN population up to very high redshifts is a key factor to understand and constrain the early growth of SMBHs, but it requires the collection of large samples of AGN. X-ray surveys, such as the 4 Ms Chandra Deep Field South (4 Ms CDFS, Xue et al. 2011), *Chandra*-COSMOS (Elvis et al. 2009), *XMM*-COSMOS (Hasinger et al. 2007) and SUBARU-*XMM* Deep Survey (SXDS; Ueda et al. 2008) are the best tool to gather such samples, since they are less biased against obscuration and galaxy dilution than optical/NIR surveys and allow a proper coverage of the $L_X - z$ plane.

Determining source obscuration is a key factor for assessing the evolution of the AGN populations. Many results at low redshift (e.g. La Franca et al. 2005; Akylas et al. 2006; Treister & Urry 2006; Hasinger 2008; Ebrero et al. 2009; Treister et al. 2009; Ueda et al. 2014) found an increasing obscured AGN fraction with increasing redshift and decreasing luminosity. However, the obscured AGN fraction at high redshift is currently unconstrained. Probing the evolution of obscuration up to high redshift would be an important tool to understand how AGN activity is triggered and how SMBH build up their mass (e.g. Menci et al. 2008). Recently, the evolution of X-ray selected high-redshift ($z > 3$) AGN population has been the topic of investigation of different authors (e.g. Brusa et al. 2009; Civano et al. 2011; Hiroi et al. 2012; Fiore et al. 2012; Lehmer et al. 2012). All of their results are consistent with a decline in the space density of luminous AGN up to $z \sim 5$, similarly to the evolution of optically selected quasars, suggesting that obscured AGN follow the same decline of unobscured AGN.

The nature of the obscuring gas is still a matter of debate. This issue is placed into a wider context known as "BH-galaxy co-evolution", which incorporates a set of recently found connections between nuclear activity and properties of the host galaxy. For instance, AGN were found to live preferentially in star-forming galaxies (e.g Rosario et al. 2012). Many authors (e.g. Rosario et al. 2013) proposed that this connection is likely a consequence of an enhanced gas content in the AGN-hosting galaxies, but so far nobody has ever properly investigated this link. The galactic gas is also one of the causes of AGN obscuration (e.g. Bianchi et al. 2012,

and references therein). Therefore, understanding the interplay between AGN and galaxies up to high-redshift is necessary to assess the processes of formation and evolution of galaxies and AGN.

This thesis is placed into this framework. We collected data from the above-mentioned X-ray surveys (and related multiwavelength coverage) to select a sample of $z > 3$ AGN with luminosity $43 \lesssim \log L_X \lesssim 45$. A spectral analysis was performed on the X-ray sources with a sufficient photon-counting statistics in order to derive parameters such as column density and luminosity. We investigated the evolution of the high-redshift AGN population by interpreting the number counts and the luminosity function we derived. We found that the space density of $z > 3$ AGN declines with redshift at least up to $z \sim 5$, refining and extending to lower fluxes and luminosities the results found by other authors. The obscured AGN fraction, where a threshold of $\log N_H = 23$ was assumed to discriminate between obscured and unobscured AGN, was estimated to be ~ 0.5. This value, when compared with low-redshift studies, suggests a positive evolution of obscuration up to $z > 3$, stronger for luminous sources.

Using FIR data in the GOODS-S, GOODS-N and COSMOS fields, we also investigated the connection between AGN and galaxy properties by deriving the gas content of the hosts. Current data forced us to restrict the study at $z < 1$, but we expect to extend it to higher redshift in the future, when ALMA data will be available. For the first time, we applied to an AGN sample a procedure to derive the gas mass from the dust mass, which is easier to estimate from the galaxy spectral energy distributions (SEDs). AGN are found to be hosted by galaxies which, on average, are much more gas-rich than inactive galaxies at similar stellar mass and redshift. This explains why they live in star-forming galaxies. The results of this thesis are advancing our understanding of the coevolution of SMBHs and galaxies: by fully measuring the cosmic history of nuclear obscuration and of the gas content in AGN hosts, we will ultimately understand the mechanisms and the physical conditions responsible for AGN triggering.

Chapter 1

Introduction

1.1 Active Galactic Nuclei and the Unification model

About 1-10% of all galaxies are characterized by very large luminosities (up to $L_{bol} \approx 10^{48}$erg/s, e.g. Wu et al. 2010), much larger than ordinary galaxies, which cannot be explained by stellar processes. This powerful emission was found to come from central and compact (< 1 pc^3) regions. The sources of such emission were called Active Galactic Nuclei (AGN). AGN spectra can be approximated as power laws over a wide range of wavelengths, suggesting a non-thermal emission mechanism.

A complex taxonomy was developed in order to classify AGN mainly on the basis of optical or radio emission characteristics. At first order, AGN are divided into Type I and Type II AGN depending on the presence or absence, respectively, of broad (up to a few 10^4 km/s) emission lines in the optical spectra, and into Radio Loud and Radio Quiet AGN on the basis of the radio emission power.

The effort to explain the different properties of the numerous classes of AGN led to the development of the so called "unified model for AGN" (e.g. Antonucci 1993, Urry & Padovani 1995, Chiaberge et al. 2000). The basic idea is that AGN belonging to different classes can be powered by the same engine, but seen in different orientations, although, in some cases, intrinsic physical diversities must be taken into account. According to that, the main elements of a typical AGN (shown in Fig. 1.1, see e.g. Peterson 2003) are:

- A supermassive black hole having mass $M = 10^{6-9} M_\odot$, the "central engine" of the AGN.
- A rotating, viscous accretion disk of gas falling towards the black hole. The temperature of the disk (up to $T \sim 10^6$ K) depends on the distance from the black hole (outer layers are cooler than inner ones) and on the black hole mass. This could explain the emission in optical/UV as a multicolor black-body, a sum of different temperature black-body spectra

$$B_\nu(T) = \frac{2h\nu^3}{c^2} \frac{1}{e^{\frac{h\nu}{kT}} - 1}. \tag{1.1}$$

  The sum of these spectra results in a power law-like spectrum $F \propto \nu^{1/3}$.

- The Broad Line Region (BLR), probably made by gaseous clouds close to the central engine (up to 100 light days) having a density $n_e \sim 10^{10-11}$ cm^{-3}, too high to emit forbidden lines, and a temperature $T \sim 10^4$ K. The emission lines (like the hydrogen Balmer series) have a large width, typically a few 10^3 km/s. The principal broadening mechanism is Doppler motion of individual clouds, being the thermal broadening $\Delta v \sim (\frac{kT}{m_p})^{0.5} \sim 10$ km/s for a $T \sim 10^4$ K gas.

- A dusty cool torus surrounding the central regions at a distance $R \sim 0.3-3$ pc. Its temperature must be lower than the dust sublimation temperature, which can vary for different dust composition (e.g. $T_{subl} \sim 900$ K for silicates, $T_{subl} \sim 1800$ K for graphite). The presence of the torus is necessary to explain why type 2 AGN are obscured, but its structure is still not well known: several models were developed proposing a homogeneous or clumpy distribution with different density and temperature profiles, or even an outflowing wind driven by a magnetic field. The torus emits thermally in IR.

- The Narrow Line Region (NLR), made by gaseous clouds more external than the BLR (up to $100 - 1000$ pc from the central engine), with a density $n_e \sim 10^{3-4}$ cm^{-3} and a temperature $T \sim 10^3$ K. The lines emitted in this region have typically a width of a few 10^2 km/s. Forbidden lines, like [OIII] at 5007 Å and 4959 Å, can be emitted under these physical conditions: the low density of the gas decreases the probability that collisional de-excitation occurs before the radiative de-excitation which, for forbidden lines, has very long time scales.

- A gaseous rarefied hot corona located above the disk having a diffuse or clumpy structure depending on the models and a temperature $T \sim 10^9$ K. The hot corona is necessary to explain the high-energy X-ray emission through Inverse Compton (IC) interaction with the low energy thermal photons emitted by the disk, but the heating mechanism is not known yet.

- Relativistic jets, perpendicular to the torus plane, which emit mostly through synchrotron mechanism from the radio band to the γ-rays and can reach the length of $R \sim 1Mpc$. The mechanism originating them is not completely understood, but it seems related to the black hole rotation-driven magnetic field (e.g. Pudritz et al. 2012). The interaction between jets and environmental medium causes the formation of the *hot spots* and the *lobes*. Only radio-loud AGN show the presence of prominent jets.

The different features of the several classes of AGN can then be explained in terms of the AGN structure and orientation, as can be seen in Fig. 1.1.

First, the presence or absence of the jets determines the main distinction between radio loud and radio-quiet AGN, respectively. Regarding radio-loud AGN, the radio emission becomes more and more important as the line of sight (l.o.s) approaches the jet axis. This is easily explained by the relativistic beaming affecting the synchrotron emission. Thus radiogalaxies, i.e. radio-loud AGN for which the l.o.s. forms a wide angle with the jet axis, are weaker radio sources than blazars, for which that angle is very small or null. The orientation does not affect the mid-IR luminosity (mainly emitted by the dusty torus itself) which indeed is equal for the two classes. Type 2 AGN (Sey 2, NLRG) are obscured by the molecular torus, hence the BLR

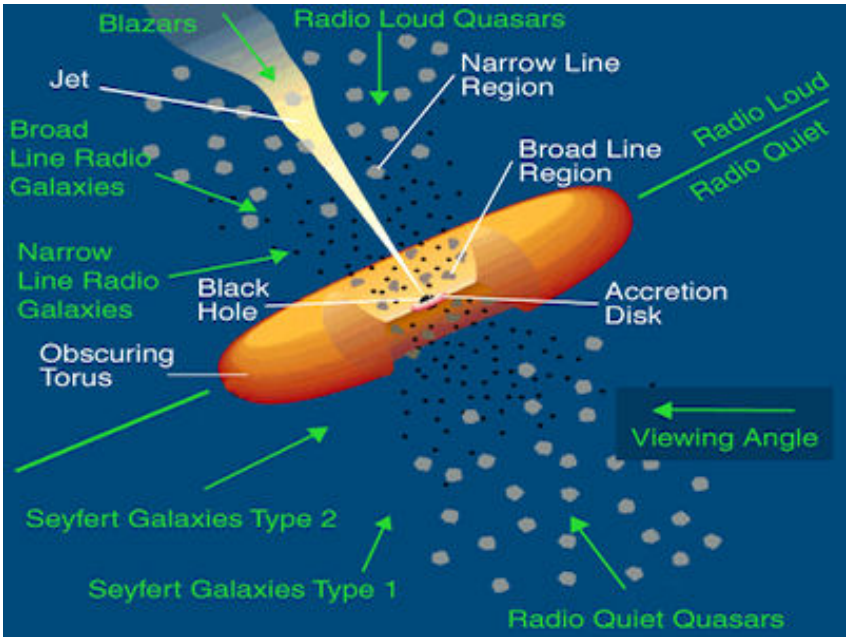

Figure 1.1: The typical AGN structure. White arrows indicate the AGN main elements according to the unified model, while green arrows refer to the l.o.s. along which the corresponding AGN class is seen. Radio loud AGN are shown in the upper part of the figure, radio-quiet AGN in the lower one. The sizes are not scaled. Adapted from Urry & Padovani (1995)

is not visible and their spectra do not show broad lines. Type 1 AGN (Sey 1, BLRG, quasars), for which the l.o.s. does not intercept the torus, show both narrow and broad lines. Blazar spectra are characterized by a very blue and featureless continuum.

1.1.1 The Black Hole paradigm

How can AGN produce such a huge amount of energy in a volume that can be estimated to be smaller than a cubic parsec? The current paradigm is the existence of a central engine consisting in a supermassive black hole surrounded by an accretion disk. The viscous material falling inwards heats up to $T \sim 10^6$ K giving rise to thermal emission. The source of energy is thus the conversion of gravitational potential energy to radiation. Assuming that the rate of energy emission is supplied by accretion:

$$L = \eta \dot{M} c^2 \tag{1.2}$$

where η is the efficiency of the process, we can find the mass accretion rate $\dot{M}$ for a typical AGN:

$$\dot{M} = \frac{L}{\eta c^2} \approx 1.8 \times 10^{-3} \frac{L_{44}}{\eta} \ \mathrm{M}_\odot/\mathrm{yr} \tag{1.3}$$

where L_{44} is the bolometric luminosity in units of 10^{44} erg/s. The potential energy of a mass particle m at a distance r from the central object of mass M is

$$U = \frac{GMm}{r} \tag{1.4}$$

hence the conversion rate of potential energy into radiation is

$$L \approx \frac{dU}{dt} = \frac{GM\dot{M}}{r} \tag{1.5}$$

From (1.2) and (1.5) we obtain $\eta \propto \frac{M}{r}$, i.e the "compactness" of the system which is maximized in case of black holes, resulting in a very high value of efficiency. We can define the size of a black hole in terms of the Schwarzschild radius R_s, i.e. the event horizon for a non-rotating black hole:

$$R_s = \frac{2GM}{c^2} \sim 3 \times 10^{13} M_8 \text{ cm} \tag{1.6}$$

where M_8 is the mass of the black hole in units of 10^8 $M_\odot$. We can estimate η in the classical limit considering a simplified case: a particle of mass m at $5R_s$. From (1.4) and (1.6) we derive

$$U = \frac{GMm}{5R_s} = 0.1mc^2 \tag{1.7}$$

Combining (1.2), (1.5) and (1.7) results in $\eta \approx 0.1$, more than an order of magnitude higher than the efficiency of thermonuclear fusion in stars (for which $\eta \approx 0.007$). Despite the simplified assumption, the efficiency obtained is a fairly good estimate and strongly suggests that the energy emitted by AGN is originated by accretion of material into a black hole. Because of the high efficiency, a quite low accretion rate can easily explain the high AGN luminosities.

Eddington Luminosity

The maximum luminosity provided by accretion of fully ionized hydrogen gas on a central object of mass M in units of solar masses $M_\odot$ assuming spherical accretion is known as *Eddington luminosity* L_E:

$$L_E \sim 1.26 \times 10^{38} \frac{M}{M_\odot} \text{ erg/s} \tag{1.8}$$

Equation (1.8) is derived considering the balance between the outward force applied by the radiation pressure $F_{rad} = \frac{L\sigma_T}{4\pi r^2 c}$ and the inward gravitational force $F_g = \frac{GMm_p}{r^2}$, where L is the luminosity of the source in *erg/s*, $\sigma_T = 6.65 \times 10^{-25}$ cm^2 is the Thomson cross-section, m_p is the proton mass and r is the distance between the proton and the central object. Under these assumptions, luminosities larger than L_E would stop the accretion. Some AGN, however, emit in a super-Eddington regime (i.e. $L > L_E$). This happens because one or more of the assumptions are wrong, particularly the spherical geometry. It can also happen during particular phases of AGN activity (probably also in very luminous high-redshift QSOs).

1.2 Continuum emission from AGN

1.2.1 Radio emission

AGN radio emission is non-thermal and it is originated by synchrotron mechanism, resulting in a power law spectrum (e.g. Risaliti & Elvis 2004, see Fig.1.2):

$$F_\nu \propto \nu^{-\alpha} \tag{1.9}$$

where F is the flux at frequence ν and α is the spectral index. Relativistic electrons having a Lorentz factor γ and a mass m_e are accelerated by a magnetic field of intensity B and emit at a characteristic frequency $\nu_S \propto BE^2$, where $E = \gamma m_e c^2$. Given an energy distribution for the electron population

$$N(E)dE \propto E^{-s}dE \tag{1.10}$$

where $N(E)$ is the number of electrons with energy in the range $[E, E + dE]$, the emission spectrum is in the form of (1.9) where $\alpha = \frac{s-1}{2}$, at least for frequencies larger than a critical (or "break") value ν_B, which depends on the medium density and the linear dimension of the emitting region. Below ν_B the spectrum is self-absorbed and $F_\nu \propto \nu^{2.5}$. Radio emission is only a very small part of the bolometric luminosity for radio-loud AGN, and is even smaller for radio quiet ones. The radio emitting regions are mainly the jets, the hot spots and the lobes.

1.2.2 IR emission

Thermal emission by the dusty torus with a temperature lower than the dust sublimation temperature, is generally invoked for IR emission (Risaliti & Elvis 2004). The dust grains absorb the UV photons emitted by the accretion disk and re-emit the radiation in the near and mid-IR bands. The spectral shape shows a maximum, the "IR bump" (see Fig.1.2), at $\lambda = 10 - 30$ μm, corresponding to a temperature $T \sim 50 - 1000$ K. This is higher by a factor of ~ 10 than the temperature of the dust heated by starlight. The peak of the AGN IR emission falls in a region of the host-galaxy SED where no other significant component is present, while in the near-IR bands a significant contribution from stars is expected and the galactic cold dust emission dominates in the far-IR. Therefore, the SED of a typical AGN shows a characteristic flattening in the mid-IR, which can be used to select AGN candidates from a parent sample of galaxies (Donley et al. 2012).

There are also hints of non-thermal IR emission for FRI radiogalaxies, which show a correlation among near IR, optical and radio nuclear luminosity suggesting the same synchrotron origin for the emission in those bands and the absence or a small contribution from the torus.

1.2.3 Optical/UV emission

The optical/UV emission is dominated by an excess of energy emitted at wavelengths shorter than 4000 Å called "Big Blue Bump" (Risaliti & Elvis 2004, see Fig.1.2), originated by multi-colour black body emission from the accretion disk, as described in § 1.1. The high energy decreasing wing of the blue bump probably

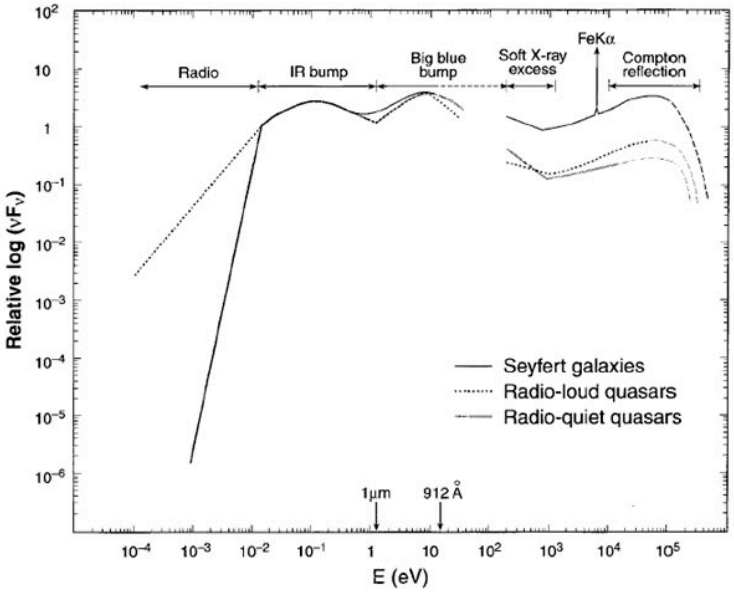

Figure 1.2: Typical broadband SED for three classes of AGN. The main features are labelled in the figure and described in the text. From Koratkar & Blaes (1999).

results in the "soft excess" below 1 keV in X-rays (red line in Fig. 1.3), although absorption by cold gas in the interstellar medium in the Milky Way prevents observation below 0.3 keV and hence does not permit to check if the two features are actually linked. Other interpretations based on a combination of reflection and absorption were also proposed to explain the soft excess (e.g. Miniutti et al. 2010). A large number of emission lines are present in optical spectra: broad permitted lines are emitted by the BLR, while narrow permitted and forbidden ones come from the NLR, as reported in § 1.1.

1.2.4 X-ray emission

X-ray emission is characterized by a main power law component (see Fig.1.2). The variability on short timescales suggests the inner region as the origin of the high energy radiation. Inverse Compton scattering (IC) is responsible for the power law emission (blue line in Fig. 1.3, see e.g. Haardt & Maraschi 1993): low-energy thermal photons emitted by the disk are up-scattered in the X-ray band by hot-thermal electrons in the corona. The resulting spectrum is a power law resulting from the superposition of many scattering orders of the intrinsic black body:

$$F_\nu \propto \nu^{-\Gamma} \tag{1.11}$$

where F_ν is in units of $photons/s/keV$, $\Gamma = \alpha + 1 \sim 1.8 - 2.0$, $\alpha = \frac{-\ln\tau}{\ln A}$ is the spectral index, τ is the optical depth, $A = e^y$ is the average energy gain in each scattering, and y is the Comptonization parameter $y = \frac{4kT}{m_e c^2}\tau$. The spectrum shows a high energy cut-off at $\sim 100 - 300$ keV, corresponding to the temperature at which electrons and photons are thermalized. The intrinsic power-law emission can be absorbed by neutral hydrogen in the proximity of the emission region (e.g. the

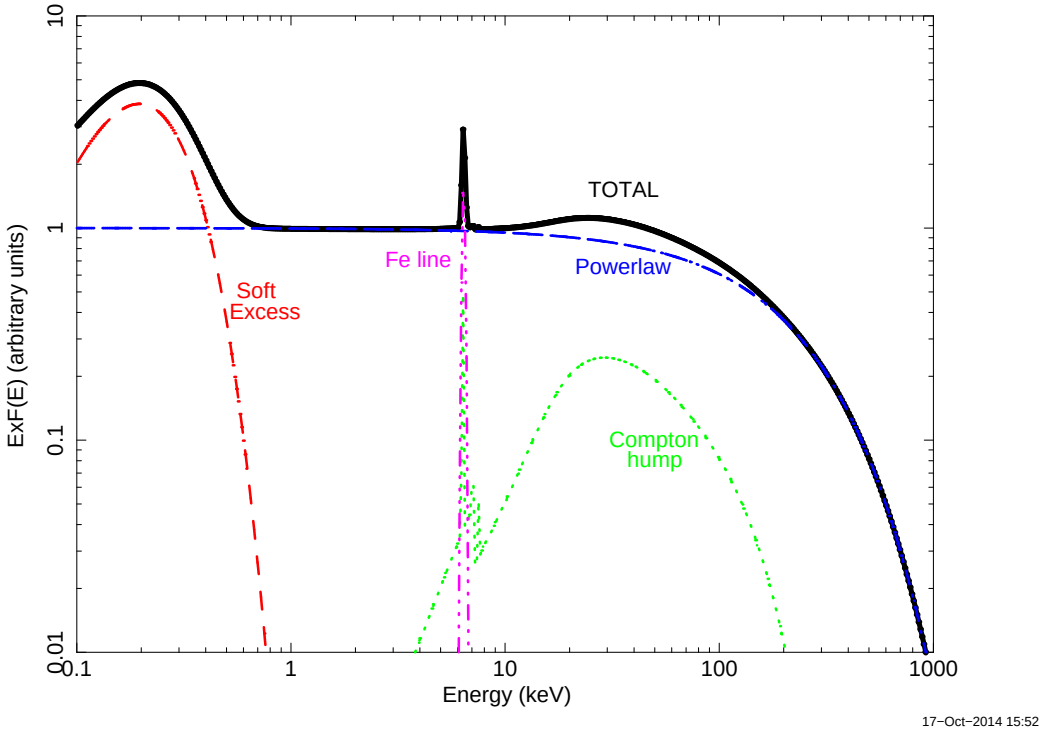

Figure 1.3: X-ray spectrum of a typical Type I AGN in arbitrary units. The black line is the total spectrum, sum of the several components shown in different colours. The red line accounts for the soft excess, possibly the high energy tail of the disk thermal emission. The blue line is the primary power law emission, with an high energy cut-off at $\sim$ 200 keV. The green line shows the reflection component. The purple line represents the Fe K_α line at 6.4 keV.

torus or gas in the host galaxy). This causes a depression of the flux at low energies and a typical photoelectric cut-off, which shift towards higher and higher energies as the column density of the absorbing material increases (Fig. 1.4). Typically, a column density of $N_H = 10^{22}\mathrm{cm}^{-2}$ is considered the threshold between unabsorbed and absorbed AGN. This subdivision roughly correspond to the optical classification in Type I and II, although the two classifications are not equivalent, an issue which is still matter of debate.

Several X-ray spectra present absorption features at $\sim$ 1 keV, due to material having an equivalent hydrogen column density of about $\sim 10^{22}$ cm^{-2}. This "warm absorber" cannot be identified with the clouds forming the BLR, the former having a higher degree of ionization. The same warm absorber could be responsible also for absorption features in UV spectrum, which moreover are blushifted relatively to the broad emission lines, leading to believe that the absorbing gas is outflowing from the central regions and possibly to identify it with the wind responsible for the absorption features in the Broad Absorption Lines (BAL) Quasars, if high velocities are involved.

The primary power-law component is reflected by the Compton thick ($N_H >> 1.5 \times 10^{24}$ cm^{-2}) disk. This reflection component (green line in Fig. 1.3) creates a

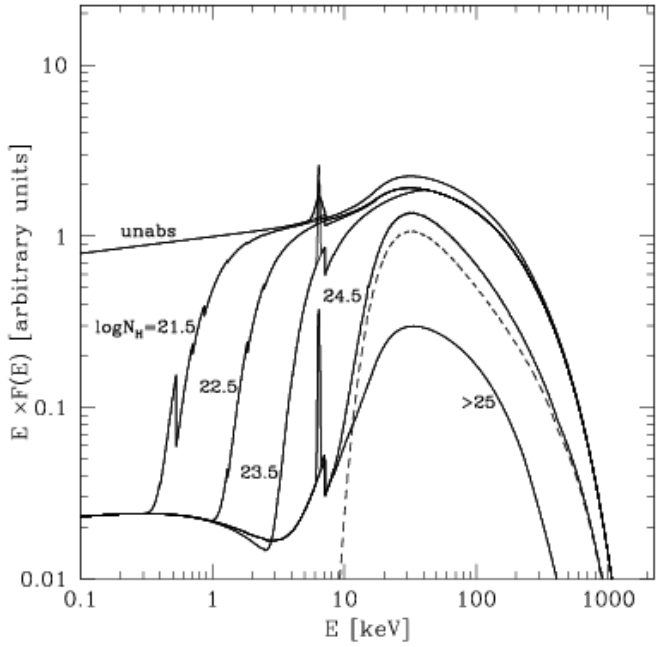

Figure 1.4: Models of AGN spectra considering different gas column density (from Gilli et al. 2007). The obscuration strongly affects the low energy part of the spectrum, resulting in a depression of the flux and in a prominent iron line, more and more evident with increasing column density.

hardening of the spectrum, the third bump (after the IR bump and the "Big Blue Bump" described in § 1.2.2 and § 1.2.3, respectively) in AGN SEDs, known as Compton or Reflection Bump, which peaks at ~ 30 keV. The intensity of this secondary component depends on many factors, such as the geometry and the ionization state of the reflecting material. Reflection from the torus can also be present.

Typical X-ray luminosities of AGN are in the range $10^42 < \log L_X < 10^46$. X-ray binaries in a galaxy can reach an overall luminosity of a few $10^{42}\,\mathrm{erg\,s^{-1}}$ in the $2-10\,\mathrm{keV}$ band (Ranalli et al. 2003). It is therefore not straightforward to discriminate between an AGN and a simple star-forming galaxy at such luminosities.

Fe K_α line

One of the most important features in X-ray spectrum is the iron K_α emission line at 6.4 keV for neutral iron (red line in Fig. 1.3) produced through fluorescence mechanism (Fabian et al. 2000) in the accretion disk and/or the dusty torus. In this process, either a K_α (6.4 keV), with a probability of $\sim 90\%$, or K_β (7.06 keV) photon is emitted, although it could be absorbed by an outer electron which afterwards would be ejected (Auger mechanism). If the Fe atom was previously ionized, there would be less electrons screening the K shell from nuclear attraction and the fluorescent photon energy would be slightly larger (up to 6.9 keV). Iron K_α is the most prominent emission line in the X-ray spectrum because Fe has the highest product among all the elements between the abundance and the fluorescent "yield"

Y, i.e. the probability of radiative de-excitation ($\sim 34\%$):

$$Y \sim \frac{Z^4}{Z^4 + 33^4} \tag{1.12}$$

where Z is the atomic number. If the Fe K_α is emitted in the innermost layers of the accretion disk, it will be affected by relativistic effects due to the presence of the black hole, in addition to the classic Doppler effect given by the disk rotation (Fabian et al. 2000). The sum of classic and relativistic transverse Doppler effect and relativistic gravitational redshift results in a red-shifted double-peaked broadened line profile, having the blue peak enhanced and a large red wing (Fig. 1.5). Several other variables, such as the inclination angle, the disk ionization state and the emissivity profile of the disk, contribute in shaping the line profile.

The strength of the Fe K_α line can be measured through its *Equivalent Width* (EW): the larger the EW, the more intense the line emission. The EW is defined as:

$$EW = \frac{1}{I_c(\nu_l)} \int I_l(\nu) d\nu \tag{1.13}$$

where $I_c(\nu_l)$ is the intensity of the continuum at the line energy and $I_l(\nu)$ is the intensity of the line.

1.3 The X-ray background and the role of obscured AGN

When an X-ray telescope is pointed towards an object of scientific interest, it collects photons from the object itself (the "signal") and photons not coming from the chosen source (the "background"). Obviously, the more the signal is "stronger" with respect to the background, the better will be the scientific output of the observation, since the derived signal information is less contaminated and more recognizable from the underlying background. The strength of the signal with respect to the background can be reported as the signal-to-noise ratio (SNR), which, in case of Poissonian statistics, as in X-ray observations is computed as

$$\mathrm{SNR} = \frac{S}{\sqrt{S + 2B}} \tag{1.14}$$

where S is the number of detected photons coming from the source of interest and B is the number of background photons, usually estimated from a nearby region of sky, free from detected sources.

The X-ray background consists of both local (e.g. due to the interactions of charged particles with the telescope material, producing fluorescent emission lines of typical elements such as Au and Ni) and cosmic (mainly unresolved sources and thermal emission from the Local Bubble) components. The cosmic X-ray background (CXB) contains a lot of information about the population of X-ray emitting objects in the Universe, since it is produced by discrete sources (Giacconi & Zamorani 1987).

The CXB spectrum presents a peak at $E \sim 30$ keV and a slope of $\Gamma \sim 1.4$ at lower energies. In order to fit it, unobscured AGN are not sufficient, since their

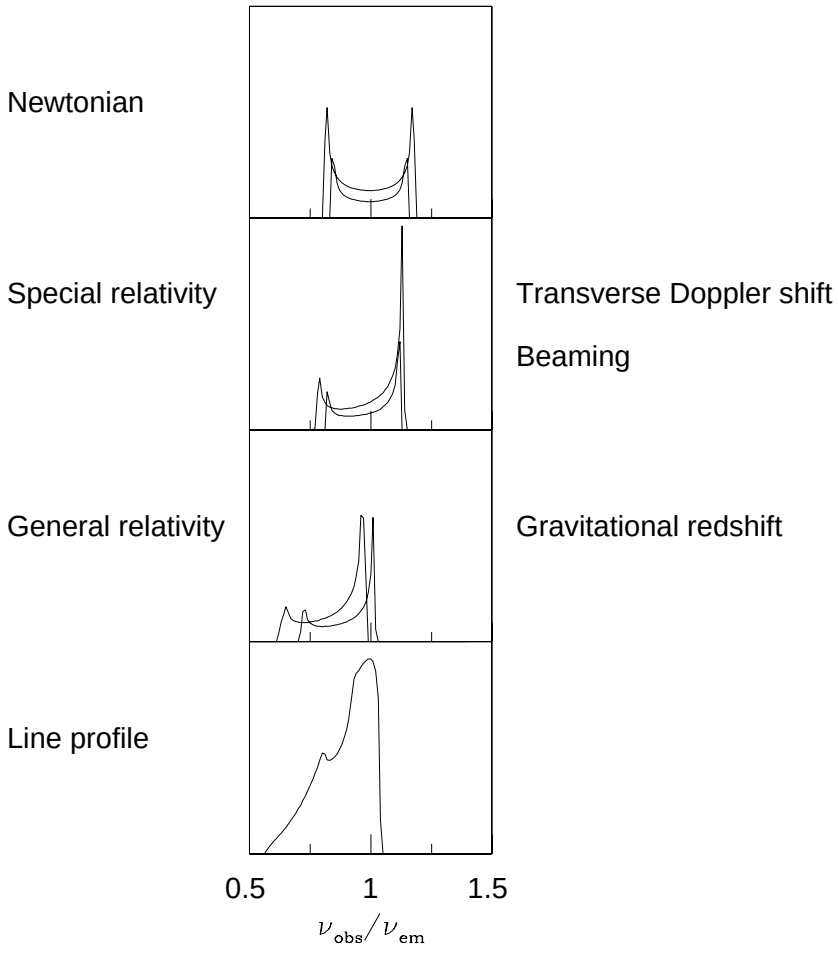

Figure 1.5: Main contributions to Fe K_α line profile (from Fabian & Miniutti 2005). From top to bottom, the classic Doppler split of a line emitted by an edge-on rotating disk, with the blue peak corresponding to the approaching part and the red one to the receding; the relativistic beaming enhancing the blue peak and the relativistic transverse Doppler effect are added; the gravitational redshift broadening the red wing and shifting the line to lower energies is added. The profiles at two different radii are represented in these panels. The last panel shows the line resulting as the sum of all the effects integrated over the entire disk.

photon indeces are too steep. Therefore, a large population of obscured AGN must be included (e.g. Setti & Woltjer 1989; Comastri et al. 1995; Gilli et al. 2001; Gilli et al. 2007).

1.3.1 Obscured AGN

As discussed in § 1.2.4, X-ray photons emitted by an AGN can be absorbed by neutral gas. If the gas column density is larger than 10^{22} cm^{-2} (i.e. the typical column densities of the galactic disks), the AGN is considered obscured. Photoelectric absorption and Compton scattering may account for the obscuration in X-rays; as an example, for a column density $N_H \sim 10^{23}$ cm^{-2}, the former dominates at $E < 3$ keV, the latter at $7 < E < 30$ keV. The incidence of absorption at $1 \leq E \leq 10$ keV becomes more and more important and the spectrum is absorbed towards higher energies as the gas column density increases (Fig. 1.4), resulting in a flattening of the observed photon index. Hence, adding obscured Compton-thin AGN contribution to that of the unobscured ones allows to match the slope of the XRB spectrum. However, even including obscured AGN in the XRB models, a discrepancy remains between observations and predictions at E $\sim$ 30 keV, as shown in the left panel of Fig. 1.6. In the context of the AGN Unification Model, the absorbing material can be identified with the torus or the BLR clouds (e.g. Bianchi et al. 2012, and references therein). Evidences for absorption from gas in the host galaxy are also found (Bianchi et al. 2012, and references therein).

X-ray observations are the best tool to detect obscured AGN, since X-rays are less biased towards obscuration with respect to optical/NIR observations (i.e. Brandt & Hasinger 2005). Several method, such as direct N_H measurement from spectral analysis (e.g. Lanzuisi et al. 2013), hardness ratio analysis (i.e. the relative number of counts detected in different bands, e.g. Brusa et al. 2009) or X-ray colours (Iwasawa et al. 2012), can be used to select obscured AGN using X-ray data. In optical/IR bands, Type II AGN can be selected, for instance, from optical spectroscopy (e.g. Mignoli et al. 2013), line diagnostic diagrams (e.g. Bongiorno et al. 2012) or colours analysis (e.g. Donley et al. 2012). Interestingly, the different classification methods applied on the same objects not always agree (e.g. Merloni et al. 2014).

The fraction of obscured AGN at $z < 3$ is found to decrease with luminosity by many works in the X-ray (Ueda et al. 2003; Steffen et al. 2003; Barger & Cowie 2005; La Franca et al. 2005; Akylas et al. 2006; Treister & Urry 2006; Della Ceca et al. 2008; Hasinger 2008; Ebrero et al. 2009; Treister et al. 2009; Brusa et al. 2010; Burlon et al. 2011; Sazonov et al. 2012; Ueda et al. 2014) and optical/IR (Simpson 2005; Treister et al. 2008; Bongiorno et al. 2010; Assef et al. 2013; Lusso et al. 2013) bands. Even though various authors have ascribed this effect to incompleteness or selection-related issues (Dwelly & Page 2006; Lawrence & Elvis 2010; Merloni et al. 2014), it could be caused by a decreasing covering factor of the obscuring material with luminosity (e.g. Maiolino et al. 2007; Wang et al. 2005; Mor & Trakhtenbrot 2011, but see also Rowan-Robinson et al. 2009). This trend can be explained in the context of a "receding torus scenario" (Lawrence 1991; Simpson 2005; Nenkova et al. 2008), in which the torus has a constant height as a function of the radius and is, at most, slowly variable as a function of luminosity. In this

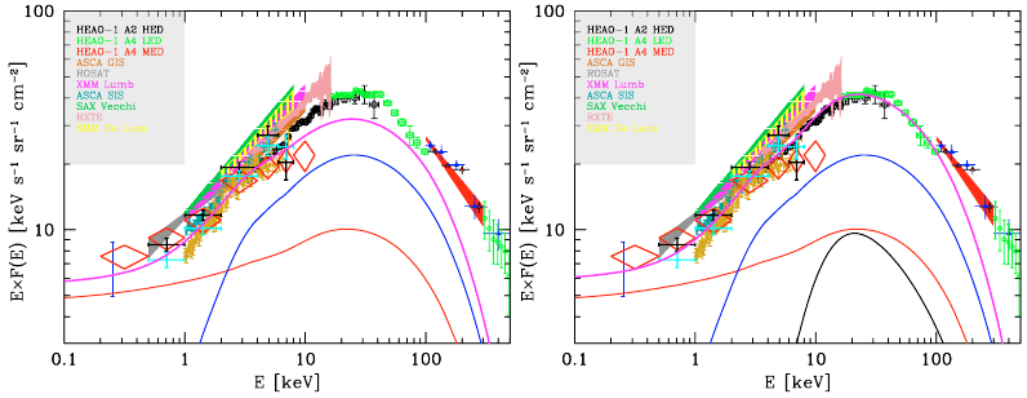

Figure 1.6: The XRB spectrum reconstructed from observations taken by different instruments and a specific parametrization of the main AGN contributions (from Gilli et al. 2007). The left panel shows the contribution of unobscured (red line) and obscured Compton-thin (blue line) AGN; in the right panel the component of obscured Compton-thick AGN (black line) is added. The resulting total spectrum (having added galaxy clusters) is displayed by the violet line in both cases. The fraction of heavily obscured AGN is required to match the observed spectrum.

case, the anti-correlation between covering factor and luminosity follows naturally from the increase of the dust sublimation radius at increasing luminosity. Another hint in favour of this behaviour is given by the so-called "X-Ray Baldwin effect" or "Iwasawa-Taniguchi effect", i.e. the decreasing equivalent width of the Fe K_α line (often considered a tracer of the covering factor) with luminosity (Iwasawa & Taniguchi 1993; Page et al. 2004; Jiang et al. 2006; Guainazzi et al. 2006; Bianchi et al. 2007). However, this scenario cannot explain the anti-correlation inferred by the X-rays studies, which are not sensitive to dust. Lamastra et al. (2006) proposed a model in which, if Compton-thin obscuration in AGN is due to the interstellar medium, the observed anti-correlation arises from the increasing gravitational pull of the SMBH at increasing mass and, hence, on statistical basis, luminosity. Alternatively, lower covering factors in luminous AGN could be due to the strong AGN radiation pressure, which would also cause massive outflows of material, as recently observed by e.g. Fischer et al. (2010); Feruglio et al. (2010); Rupke & Veilleux (2011); Sturm et al. (2011); Cano-Díaz et al. (2012).

1.3.2 The role of Compton-Thick AGN

For column densities exceeding the inverse of Thomson cross-section $N_H \geq \sigma_T^{-1} \sim 1.5 \times 10^{24}$ cm^{-2}, the AGN is defined as Compton thick (CT). In this case the spectrum is largely depressed up to $E \sim 10$ keV for mildly CT sources ($24 \lesssim \log\frac{N_H}{cm^{-2}} \lesssim 25$), while almost the entire high-energy radiation is shielded for heavily CT AGN ($\log\frac{N_H}{cm^{-2}} \gtrsim 25$). The main features in the X-ray spectrum of a CT AGN are the strong Fe line ($EW \sim 1$ keV, since the continuum is depressed by strong absorption) and the peculiar spectral shape at $E = 20-30$ keV, due to reflection probably from the inner side of the obscuring torus. The extremely hard photon index of CT sources is necessary to cover the gap in the XRB spectrum, as can be seen in the

right panel of Fig. 1.6. Models predict that CT AGN (mainly at $z \sim 1$) contribute from 10% to 30% to the XRB peak at 30 keV. Since they are fainter than obscured and unobscured AGN, their intrinsic fraction is therefore larger than these values.

Despite of the large number predicted (e.g. Gilli et al. 2007) and estimated (e.g. Vignali et al. 2010; 2014), only some tens of CT AGN are known in the local universe (e.g. Della Ceca et al. 2008, and references therein) since they are generally less luminous than the less absorbed ones and their X-ray emission can be best revealed at very high energies ($E > 10$ keV), a spectral range difficult to observe with current instruments. The most efficient methods to identify heavily obscured sources are based on the X-ray spectral properties (i.e. by revealing a strong iron line over a flat reflection continuum) and IR observation, even though there can be biases from dusty galaxies using the latter method. Larger sample of sources and deep surveys are required to better understand the physics and evolution of CT AGN, especially at high redshift, where just a handful of them are known (e.g. Norman et al. 2002; Comastri et al. 2011; Gilli et al. 2011; 2014).

1.4 Evolution of the X-ray selected AGN population

The AGN comoving space density is not constant over time: there are much more AGN at $z > 1$ than in the local universe. Many works were performed in order to understand the evolution of the AGN population from the Local Universe to high redshift. However, there still are many open questions. In particular, the behaviour at high redshift ($z > 3$) is not well constrained, implying severe uncertainties in AGN evolution and in the interplay between AGN and host galaxies at the early ages of the Universe. Moreover, the density and the role of heavily obscured AGN are mostly unconstrained beyond the Local Universe. Deep X-ray surveys are needed to resolve those problems by allowing us to collect large samples of AGN, especially at high redshift. In this section, we summarize the current knowledge of the evolution of the AGN population from low to high redshift and some of the open issues we want to address in the following chapters.

1.4.1 Low-redshift evolution

The evolution of the AGN space density at $z < 3$ depends on the luminosity (see Fig. 1.7), with the space density of low-to-moderate luminosity ($L_X < 10^{44} \mathrm{erg\,s^{-1}}$) peaking at lower redshift ($z \sim 1$) than that of high-luminosity ($L_X > 10^{44} \mathrm{erg\,s^{-1}}$) AGN ($z \sim 2$). This effect is known as "AGN downsizing" and is similar to the cosmic downsizing of galaxies (e.g. Cowie et al. 1996; Fontanot et al. 2009, and references therein). Several analytical models were proposed in order to parametrize this behaviour. The two most popular models are the Luminosity Dependent Density Evolution model (LDDE; Miyaji et al. 2000; La Franca et al. 2002; 2005; Ueda et al. 2003; 2014; Hasinger et al. 2005; Silverman et al. 2008; Ebrero et al. 2009), originally proposed by Schmidt & Green (1983) to describe the evolution of optically-selected AGN, in which the space density of AGN evolves differently for different luminosities and there is no actual evolution in luminosity, and the Luminosity And Density Evolution (LADE) model (Aird et al. 2010), which assumes that the AGN population evolves independently in luminosity and density with cosmic time.

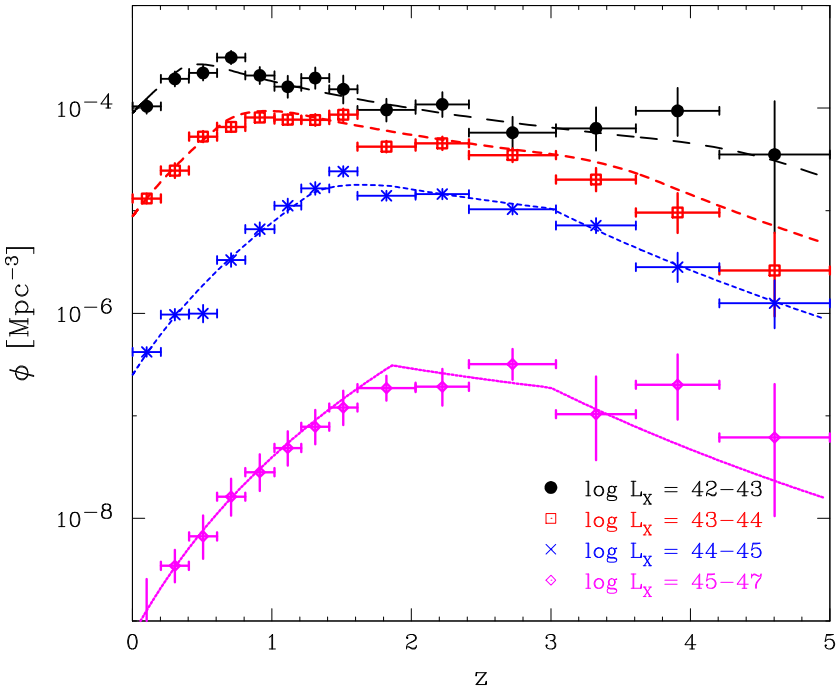

Figure 1.7: AGN space density as a function of redshift, for 4 different X-ray luminosity bins (from Ueda et al. 2014). The "AGN downsizing" is evident from the peak of the space density shifting at higher redshift for increasing luminosity.

Beside the evolution of the whole AGN population, it is important to assess the evolution of the obscured AGN fraction, usually defined as the fraction of AGN obscured by a column density $N_H > 10^{22} cm^{-2}$ with respect to the total number of AGN. The obscured AGN fraction is found to increase at least up to $z = 2 - 3$ (La Franca et al. 2005; Akylas et al. 2006; Dwelly & Page 2006; Tozzi et al. 2006; Treister & Urry 2006; Hasinger 2008; Ebrero et al. 2009; Treister et al. 2009; Merloni et al. 2014; Ueda et al. 2014), even though other authors found no evidence in favour of such an evolution (Ueda et al. 2003; Lusso et al. 2013).

A possible evolutionary scenario emerging from recent works (e.g. Hopkins et al. 2008; Hickox et al. 2009) suggests that the activity of luminous ($L_{2-10\ keV} \geq 10^{44}$ erg s^{-1}) AGN, triggered by major gas-rich mergers (e.g. Di Matteo et al. 2005; Menci et al. 2008), occurs in short bursts (with timescales of the order of ~ 0.01 Gyr; e.g. Alexander et al. 2005). In this case, the gas accretion is chaotic (Hopkins et al. 2008) and may produce a large covering factor and column densities. This phase of strong accretion is hidden by the Compton Thick material, until it is swept away by the emitted radiation and the AGN shines as an unobscured quasar. Since the merging rate (Kartaltepe et al. 2007) and, most importantly, the gas fraction (e.g. Carilli & Walter 2013) increases with redshift, the fraction of obscured AGN may then increase as well. At lower luminosities, the accretion onto the SMBH is probably driven by secular processes (~ 1 Gyr; e.g. Daddi et al. 2007; Elbaz et al. 2011) and the accretion is expected to be smoother and symmetrical. Therefore, the fraction of obscured AGN would be quite constant with

redshift, given that obscuration would be a function of the system geometry only. This luminosity-dependent evolutionary path is therefore partly in contrast with the traditional Unification model, where the different AGN properties are simply due to orientation effects. In that case, the evolution of the obscured AGN fraction is expected not to depend on the luminosity.

1.4.2 High-redshift ($z > 3$) evolution

Assessing the evolution of the AGN population, especially at high redshift, when galaxies formed, is essential to understand the galaxy formation and evolution, in which AGN are fundamental actors. Moreover, it is necessary to constrain the models of BH seeds formation in the initial phases of the Universe. AGN could also play an important role in the re-ionization of the Universe at $z \gtrsim 6$, which could be quantified only if we knew the properties of high-redshift AGN and how they evolved. However, the evolution of the AGN population, especially at low luminosities ($L_X < 10^{44}\,erg\,s^{-1}$) at $z > 3$ is still mostly unconstrained. That is because of the difficulty in collecting large, complete and reliable samples of $z > 3$ AGN. Therefore, works aimed at studying the AGN population over a wide range of redshift preferentially constrain the evolution at low redshift. This is the case, for instance, of Aird et al. (2010), who found that a LADE model well describes the evolution of the AGN population from $z = 0$ to $z \sim 3.5$. Recently, Ueda et al. (2014) proposed a LDDE model with two different critical redshift at which the density evolution changes slope. They also suggested that AGN at high-redshift underwent an "upsizing" evolution, in the sense that the ratio of the space density of low and high-luminosity AGN increases at increasing redshift. However, all the parameters affecting the luminosity function evolution at $z > 3$ were fixed to specific values (e.g. they used the slope by Hiroi et al. 2012 to characterise the decline of the AGN space density up to $z \sim 5$). Their results are therefore strongly driven by the low-redshift behaviour.

Recently, deep X-ray data has made possible for many author to focus on the $z > 3$ evolution. Brusa et al. (2009) selected a sample of 40 $3 < z < 4.5$ X-ray selected AGN from the *XMM*-COSMOS survey. They found the space density to decline with redshift similarly to optically-selected QSOs (e.g. Schmidt et al. 1995). Civano et al. (2011) confirmed this behaviour by studying a sample of 81 $3 < z < 5.3$ AGN from the *Chandra*-COSMOS survey. The same conclusion was reached by Lehmer et al. (2012) by analysing the number counts of high-redshift AGN in the 4 Ms *Chandra* Deep Field South. Hiroi et al. (2012) expanded the Civano et al. (2011) work by adding data from the SUBARU/*XMM* Deep Survey and found that the space density declines as $\Phi \propto (1 + z)^{-6.2}$ (see Fig. 1.8). Kalfountzou et al. (2014) derived consistent results from the largest (209 objects) sample of $z > 3$, X-ray selected AGN to date, using data from the *Chandra*-COSMOS and ChAMP (Kim et al. 2007; Green et al. 2009) surveys.

However, we do not know much about the space density of $L_X < 10^{44}\,erg\,s^{-1}$ AGN at high redshift. In particular, we do not know if they follow the same decline of high-luminosity AGN, or if their evolution is different, as at lower redshift. For instance, Ueda et al. (2014) suggested that at $z > 3$ the space density of low-luminosity AGN declined at a slower rate with redshift than that of QSOs-like

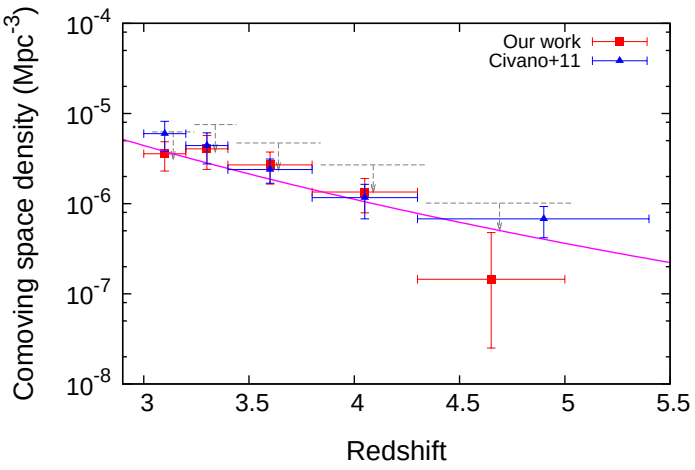

Figure 1.8: AGN space density as a function of redshift. Red squares and grey downward-pointing arrows refer to the measurements and upper limits (in the most conservative case taken into account), respectively, of the space density of $\log L_X > 44$ AGN from Hiroi et al. (2012). Blue triangles are the results from Civano et al. (2011) at $\log L_X > 44.15$. The magenta curve is the best-fit model in the form $\Phi \propto (1+z)^{-6.2}$, computed by simultaneously fitting the two data-set. From Hiroi et al. (2012).

objects (see Fig. 1.7).

Another open issue regards the obscured AGN fraction at high redshift. Even though X-rays are less biased towards obscuration than optical/NIR bands and despite the positive K-correction favouring the detection of obscured sources at high redshift, obscured AGN are difficult to be detected at high redshift, since they are typically fainter and hard to be identified in optical surveys than unobscured objects. Therefore, computing the obscured AGN fraction at high redshift, where the available samples of AGN are small, is not an easy task. Hiroi et al. (2012) reported an obscured ($N_H > 10^{22} cm^{-2}$) AGN fraction of $0.54^{+0.17}_{-0.19}$ in the redshift range $3 < z < 5$ and luminosity $\log L_X = 44 - 45$, a factor ~ 2.5 higher than the estimate by Burlon et al. (2011) in the Local Universe. Larger samples are needed to derive stricter constraints. Moreover, there is no estimate on the fraction of obscured, low-luminosity AGN at high redshift. The assessment of the obscured AGN fraction at high redshift would be crucial to constrain possible different evolutionary path for AGN of different luminosity, as described in § 1.4.1.

1.4.3 BH seeds

Recently, SMBH with masses $\gtrsim 10^9$ $M_\odot$ were discovered at $z > 6$ (Willott et al. 2003; Fan et al. 2006b;a; Willott et al. 2009; Mortlock et al. 2011), when the Universe was less than 1 Gyr old. Models of SMBH formation and growth must face the challenge represented by the presence of these massive objects when the Universe was so young. Eddington-limited or even super-Eddington accretion is usually required

by the models of BH seed formation to reach such masses in a few hundreds million years. The two most promising models involve light or heavy seeds. Light seeds are usually associated to remnants of POP III stars with masses $\sim 10^2$ $M_\odot$ (Madau & Rees 2001; Alvarez et al. 2009; Volonteri & Begelman 2010, see also Whalen & Fryer 2012). Models involving heavy seeds predict SMBH seeds to be formed by monolithic collapse of massive pristine gas clouds, possibly passed through a phase of "quasi-star" (Haehnelt & Rees 1993; Loeb & Rasio 1994; Eisenstein & Loeb 1995; Bromm & Loeb 2003; Koushiappas et al. 2004; Begelman et al. 2006; Lodato & Natarajan 2006; Begelman 2010). Other models deal with stellar dynamics in star clusters (Devecchi & Volonteri 2009). Assessing the evolution of the AGN population up to very high redshift can put constraints on the different models of SMBH seed formation.

1.5 AGN and host galaxies

Though the sphere of influence of the SMBH which reside in the centre of almost all galaxies is just a tiny fraction of the volume of the host galaxies themselves, their presence is thought to influence the galactic environment, especially during the active phases. In this section, we discuss some of the relations found between AGN and host-galaxy properties, which can be placed in a scenario in which the evolution of BH and galaxies are strongly coupled.

1.5.1 BH-galaxy co-evolution

Several authors have reported a correlation between the mass of the central BH and the properties of the host galaxy, such as the stellar velocity dispersion in the bulge, the mass and luminosity of the bulge (e.g. Magorrian et al. 1998; Ferrarese & Merritt 2000; Marconi & Hunt 2003; Häring & Rix 2004; Bennert et al. 2011; McConnell & Ma 2013; Kormendy & Ho 2013, and references therein; see Fig. 1.9). These relations imply that the growth of the SMBH and the evolution of the host galaxies are somehow connected. Hints in favour of a coupled evolution come also from the very similar evolution of the BH accretion rate (BHAR; e.g. Boyle & Terlevich 1998; Marconi et al. 2004; Aird et al. 2010; Delvecchio et al. 2014) and star formation rate (SFR; e.g. Hopkins & Beacom 2006) densities over cosmic time. Both of these quantities peak at $z \sim 2$ and then decrease towards higher and lower redshifts (see Fig. 1.10). Usually, the scenario in which BH and galaxy evolutions are coupled is referred to as "BH-galaxy co-evolution". However, several key issues remain still open. In particular, it is still unclear if the co-evolution is the result of a direct interaction between BH and host galaxy, possibly driven by feedback processes due to the massive outflows ejected during the accretion phase seen in many AGN (e.g. Feruglio et al. 2010; Maiolino et al. 2012; Cicone et al. 2014; Harrison et al. 2014), or if a third quantity is in play and is responsible of the similar evolution of BH and galaxies (e.g. Jahnke & Macciò 2011).

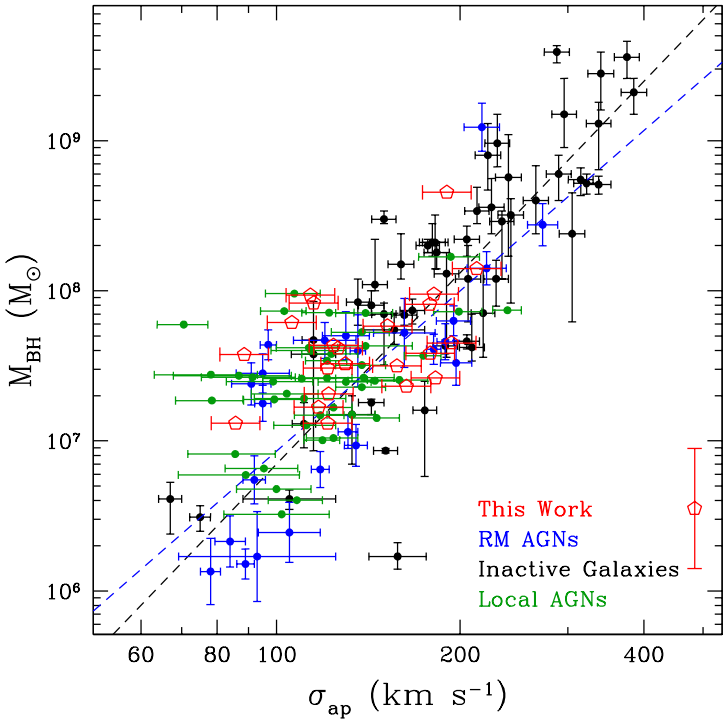

Figure 1.9: Collection of BH masses plotted against the stellar velocity dispersion of the bulge of the host galaxies (from Bennert et al. 2011).

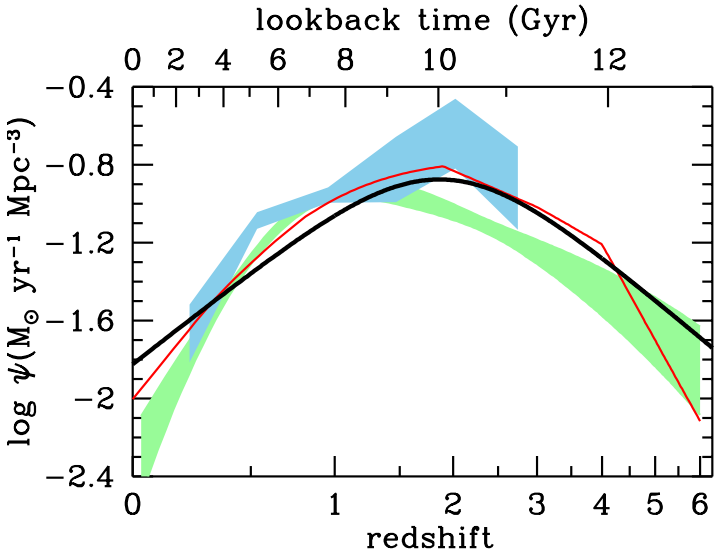

Figure 1.10: Best estimate of the star-formation rate density (black line), overplotted to the black-hole accretion rate density estimates (scaled up by a factor of 3300) by Shankar et al. (2009, red curve), Aird et al. (2010, green stripe) and Delvecchio et al. (2014, blue stripe). From Madau & Dickinson (2014).

1.5.2 AGN and star formation

The similar evolution of the BHAR and SFR densities suggests that a relation could be in place between the presence and properties of AGN, which trace the accretion phases in which SMBH grow in mass, and the star formation properties of the host galaxies. A correlation between AGN luminosity and SFR was found at least for QSO-like objects (e.g. Lutz et al. 2008; 2010; Shao et al. 2010; Rosario et al. 2012). More controversial results were found for low-luminosity AGN, but many authors reported that AGN prefer to live in star-forming galaxies (e.g. Silverman et al. 2009; Santini et al. 2012b; Rosario et al. 2013). However, this link is quite weak and not well established, especially at low luminosity.

The gas content is often regarded as a more fundamental property of galaxies, with respect to the SFR. The SFR is tightly related to the gas content through the Schmidt-Kennicutt relation (Schmidt 1959; Kennicutt 1998, SK relation hereafter). Currently, one of the most favoured scenarios is that the cosmic evolution of the star formation rate in galaxies is mostly a consequence, through the SK relation, of the more fundamental evolution of their (molecular) gas content (e.g. Obreschkow & Rawlings 2009; Lagos et al. 2011). Within the context of AGN, gas is the fundamental ingredient both for nuclear activity and star formation. Moreover, gas in AGN-hosting galaxies can play an important role in absorbing the nuclear emission. Possible differences in terms of star formation properties between AGN hosts and inactive galaxies could be due to more fundamental differences in terms of gas content, as recently argued by Santini et al. (2012b) and Rosario et al. (2012; 2013). Therefore, it is most important to obtain information on the gas content of AGN host galaxies, possibly as a function of galaxy properties (e.g. stellar mass) and redshift.

Chapter 2

The high-redshift ($z > 3$) AGN population in the 4 Ms *Chandra* Deep Field South

In this chapter, we present results from a spectral analysis and the number counts of a sample of high-redshift ($z > 3$) X-ray selected AGN in the 4 Ms *Chandra* Deep Field South (CDFS). In particular, we have selected a sample of "bona fide" high-redshift ($z > 3$) AGN using the most recent spectroscopic and photometric information available, and performed an X-ray spectral study to understand their properties and, possibly, constrain the early phases of AGN evolution. X-rays are less biased towards obscuration and galaxy dilution with respect to optical/NIR bands, allowing us to select large samples of objects and not to lose important sub-populations of AGN (e.g. heavily obscured objects). Moreover, high-redshift objects detected in the X-rays, at the sensitivities of current surveys, can be directly considered *bona fide* AGN, resulting in virtually uncontaminated samples of AGN. Being the deepest X-ray survey to date (see § 2.1), the 4 Ms CDFS provides the best data to study the spectral properties, like obscuration, of Seyfert-like AGN at high redshift and the number counts down to faint fluxes, where it is still not well constrained. We will also compare the results with previous findings in literature and predictions from semi-analytical models. In § 2.1 we describe the main characteristics of the 4 Ms CDFS, in § 2.2 we describe the selection of a sample of $z > 3$ AGN, in § 2.3 we report the spectral analysis procedure, in § 2.3.3 we study the column density distribution of the sample, in § 2.3.4 we derive the number counts, in § 2.4 we discuss the results. This chapter is then briefly summarised in § 2.5. Throughout this work we assume a $\Omega_m = 0.27$, $\Omega_\Lambda = 0.73$ and $H_0 = 70$ $\mathrm{km\,s^{-1}Mpc^{-1}}$ Universe (Spergel et al. 2003). The errors on the spectral parameters correspond to the 90 per cent confidence level for one parameter of interest (Avni 1976). All the results are presented in Vito et al. (2013).

2.1 4 Ms CDFS

The 4 Ms *Chandra* Deep Field South (CDFS; Xue et al. 2011) is the deepest X-ray survey to date. It consists of 54 observations for a total of 3.872 Ms taken with

the Advanced CCD Imaging Spectrometer imaging array (ACIS-I), on board of the *Chandra* X-ray satellite: the first 23 observations, corresponding to ≈ 2 Ms, were taken between October 1999 and November 2007 (Giacconi et al. 2002; Luo et al. 2008), while the second block between March 2010 and July 2010 (Xue et al. 2011). The aim points and roll angles are different for individual observations, resulting in a total covered area of ≈ 464.5 arcmin^2. The average aim point is $\alpha = 03^h32^m28.06^s$ $\delta = -27°48'26.4''$, where α and δ are the J2000 right ascension and declination. The focal plane temperature was $-120°$ during all observations but the first two, for which it was $-110°$. We chose to exclude the two $-110°$ observations during this work because they need different data reduction procedures. We used the information about the X-ray sources catalogue, photometry, exposure maps and redshift provided in Xue et al. (2011).

Fig. 2.1 shows the full band (0.5-8.0 keV) exposure map of the 4 Ms CDFS, corrected for vignetting (i.e. the fractional loss of photons at increasing off-axis angle). The effective exposure time drops in peripheral areas because the vignetting effect increases with the off-axis angle and less observations overlap in the external regions of the field, while almost all observations overlap in the internal ones. The 4 Ms CDFS on-axis sensitivity limit is $\approx 3.2\times10^{-17}/9.1\times10^{-18}/5.5\times10^{-17}$ ergs cm^{-2}s^{-1} in the $0.5-8.0/0.5-2.0/2.0-8.0$ keV band (Xue et al. 2011), about a factor 2 of improvement with respect to the 2 Ms CDFS. The sensitivity is a function of the off-axis angle: the highest sensitivity is reached on-axis (Fig. 2.2).

The 4 Ms CDF-S main catalog (Xue et al. 2011) contains 740 X-ray point sources (AGN, normal galaxies and stars), 634/650/403 of which were detected in the full/soft/hard ($0.5-8/0.5-2/2-8$ keV) band. Fig. 2.3 shows the composition of three 4 Ms CDF-S images in different bands.

Multi-wavelength identification was carried on using optical, near-infrared and radio data provided by Giavalisco et al. (2004),Grazian et al. (2006),Caldwell et al. (2008),Miller et al. (2008),Taylor et al. (2009), Damen et al. (2011).

Out of the 740 sources in the main catalog, 716 ($\approx 96.8\%$) have multi-wavelength counterpart. This high identification rate, combined with the small false-matching rate ($\approx 5.9\%$), is an independent measurement of the reliability of X-ray sources in the catalog.

Spectroscopic redshifts were collected Le Fèvre et al. (2004), Szokoly et al. (2004), Mignoli et al. (2005), Ravikumar et al. (2007), Vanzella et al. (2008), Popesso et al. (2009), Treister et al. (2009), Balestra et al. (2010) and Silverman et al. (2010), while photometric redshifts, as well as their 1σ upper and lower bounds, were collected from Luo et al. (2010), which also provided an alternative secondary photometric redshift solution, Cardamone et al. (2010) and Rafferty et al. (2011). Among the 716 sources with multi-wavelength counterpart, 419 (58.5%) have spectroscopic redshift measurement and 668 (93.3%) have photometric-redshift estimates from at least one of the three considered papers.

A source is identified by Xue et al. (2011) as an AGN if at least one of the following 5 criteria is satisfied:

1. Intrinsic X-ray luminosity $L_{0.5-2\ \mathrm{keV}} \geq 3\times10^{42}$ erg s^{-1}. The estimate of the luminosity is provided by the main catalog itself.

2. Effective photon index, derived from the counts in the hard and soft band, of

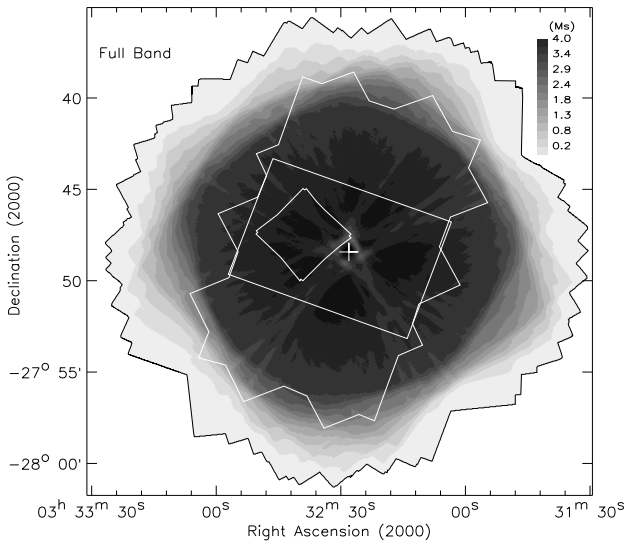

Figure 2.1: Full band (0.5-8.0 keV) exposure map of the 4 Ms CDFS (from Xue et al. 2011). The weighted average aim point is indicated with a plus sign. The large polygon, rectangle and square represent the GOODS-S (Giavalisco et al. 2004), the CANDELS GOODS-S (Guo et al. 2013) and the *Hubble* Ultra Deep Field (Beckwith et al. 2006), respectively.

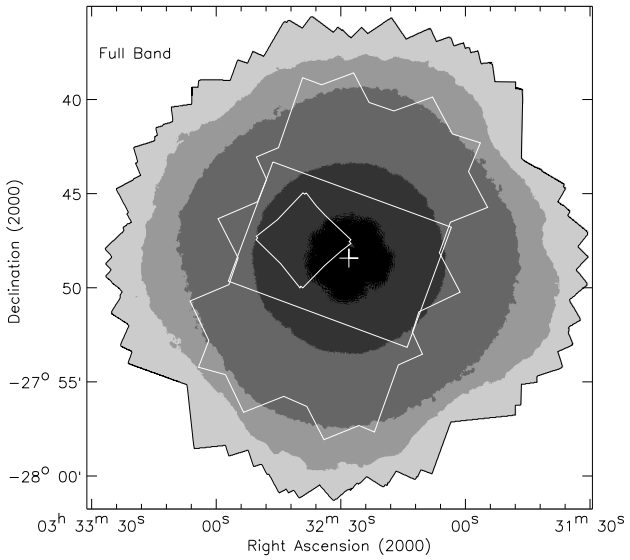

Figure 2.2: 4 Ms CDF-S full band sensitivity map (from Xue et al. 2011). The different colors represent areas for which the flux limit is $< 4.0^{-17}$, 4.0^{-17} to 10^{-16}, 10^{-16} to 3.3×10^{-16}, 3.3×10^{-16} to 10^{-15}, $> 10^{-15}$ ergs cm^{-2}s^{-1} from the darker to the lighter one. The regions and the plus sign are the same as in Fig. 2.1

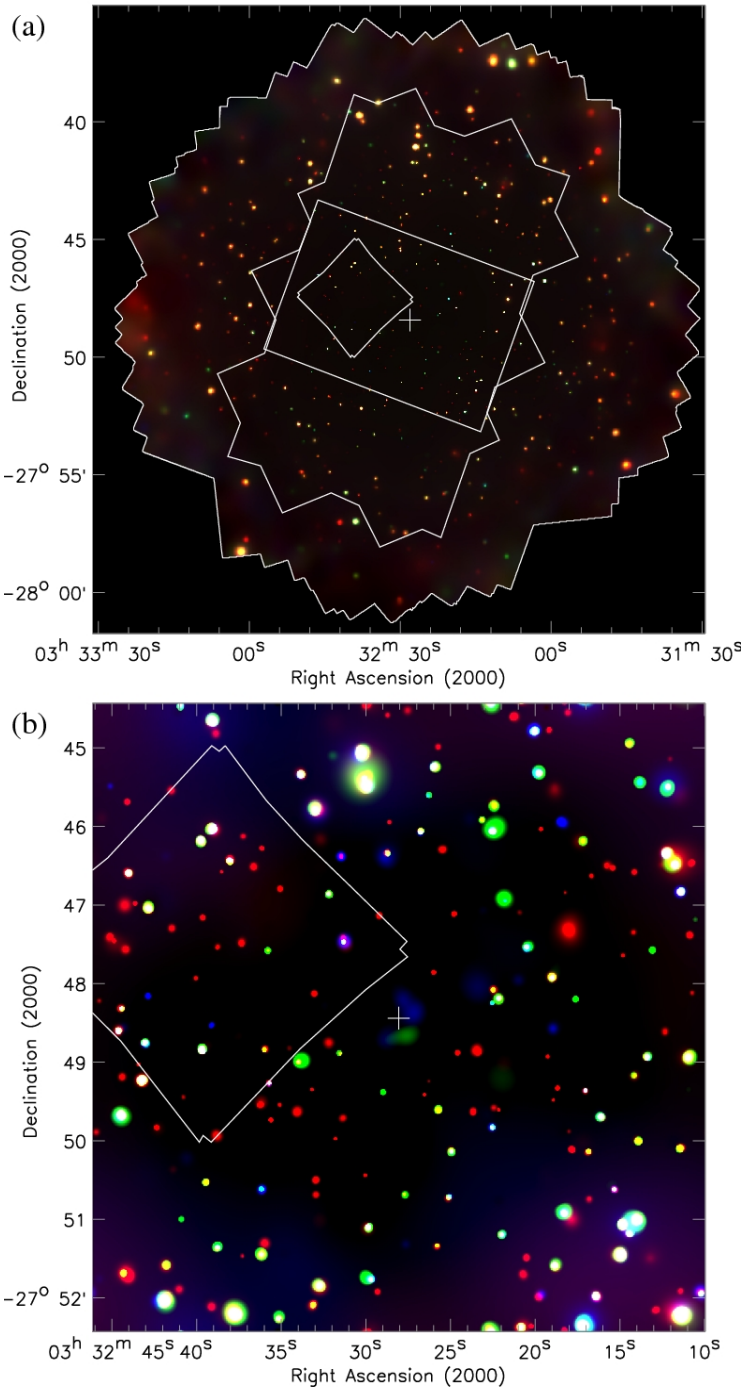

Figure 2.3: (a) False-color 4 Ms CDF-S image, obtained overlapping exposure corrected smoothed images in three bands: $0.5-2.0$ keV (red), $2.0-4.0$ keV (green) and $4.0-8.0$ keV (blue). (b) A zoom in the central $8'\times8'$ region (a different contrast ratio is used). The regions and the plus are the same as in Fig. 2.1. From Xue et al. (2011).

$\Gamma \leq 1.0$. This criterion mainly individuate obscured AGN.

3. X-ray-to-optical flux ratio of $\log(f_X/f_R) \geq -1$, where the X-ray flux is in the soft or hard or full band and is computed using the net counts.

4. X-ray emission exceeding the level expected from pure star formation.

5. AGN features (e.g. broad emission lines and/or high-excitation emission lines) detected in the optical spectrum.

AGN represent 76.8% (568 out of 740) of the main-catalog sources and reach a surface density of 9800^{+1300}_{-1100} deg^{-2} within an off-axis angle of $\theta \leq 3'$.

2.2 The high-redshift sample

We assembled a sample of X-ray selected AGN at $z > 3$ in the 4 Ms CDF-S on the basis of both spectroscopic and photometric redshifts. Spectroscopic redshifts for $z > 3$ AGN were collected from Szokoly et al. (2004), Vanzella et al. (2008), Popesso et al. (2009), Silverman et al. (2010), Vanzella et al. (2010), Wuyts et al. (2009) and E. Vanzella (private communication[1]). A quality flag ("secure/insecure") is associated to them in Xue et al. (2011), except for spectroscopic redshifts from Vanzella et al. (2010), Wuyts et al. (2009) and E. Vanzella (private communication), which are not used by Xue et al. (2011) and for which we retrieved the original quality information. Photometric redshifts were gathered from Luo et al. (2010, which also provides secondary solutions), Cardamone et al. (2010), Rafferty et al. (2011), Santini et al. (2009), Wuyts et al. (2008), Taylor et al. (2009) and Dahlen et al. (2010) .

We first cross-matched the X-ray positions of all sources in the 4 Ms CDF-S main catalogue (740 sources; Xue et al. 2011) with the positions of the sources in the catalogues which are not collected by Xue et al. (2011) using a 1 arcsec matching radius. Considering the other catalogues, we assumed the counterparts reported by Xue et al. (2011), which used a likelihood-ratio matching technique. We pre-selected a sample of high-redshift AGN candidates, including all sources which have a spectroscopic or photometric redshift $z > 3$ in at least one of the considered catalogues (96 sources). Given the large amount of available information, we built up a set of criteria to select the final high-redshift sample. A source is included in the sample if it fulfils one of the following requirements:

(1). A spectroscopic redshift flagged as "secure" $z_{spec,s} > 3$ is provided.

(2). A spectroscopic redshift flagged as "insecure" $z_{spec,i} > 3$ is available and it is upgraded to "secure" by re-analysing the optical spectrum, checking the quality and reliability of the identified spectral features.

(3). $N_{phot,>3} - N_{phot,<3} \geq 3$, where $N_{phot,>3}$ is the number of photometric redshifts larger than 3 and $N_{phot,<3}$ is the number of photometric redshifts lower than 3, when no secure spectroscopic redshift and more than two photometric redshifts are available.

[1]These spectroscopic redshifts are based on a re-analysis of VIsible MultiObject Spectrograph (VIMOS) and FOcal Reducer and low dispersion Spectrograph (FORS2) spectra.

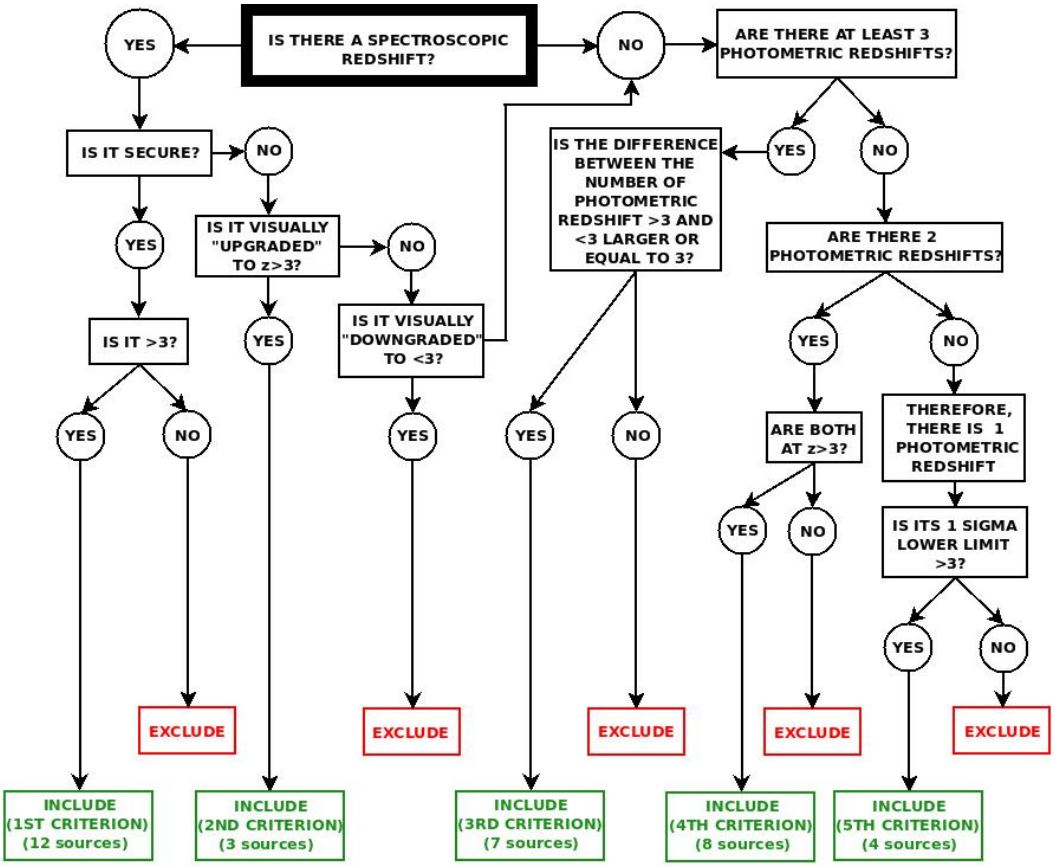

Figure 2.4: Selection criteria applied to the 96 pre-selected sources (which have at least one spectroscopic or photometric redshift $z > 3$) to be included in the high-redshift sample, visualised as a flow chart. The starting block has bold sides.

(4). When no secure spectroscopic redshift and two photometric redshifts are available, both photometric redshifts are larger than 3.

(5). When no secure spectroscopic redshift and only one photometric redshift is available, its 1σ lower bound (provided by the original catalogue) is larger than 3.

The selection procedure is shown in Fig. 2.4 as a flow chart. We will discuss in § 2.4 the effects of relaxing the selection criteria on the sample size.

The final sample consists of 34 sources (see Tab. 2.1). A spectroscopic redshift is adopted for 15 sources (12 and 3 are included in the high-redshift sample on the basis of the first and second criterion, respectively) and a photometric redshift is adopted for 19 sources (7/8/4 sources are included in the sample on the basis of the third/fourth/fifth criterion, respectively). XID 392 has no redshift information in Xue et al. (2011) and a photometric redshift $z = 6.22$ from Santini et al. (2009), but no error on the redshift is provided. Therefore, we conservatively excluded this source from the sample. The sample presented in this work is a factor of ~ 2 smaller than that discussed in Lehmer et al. (2012). This difference will be discussed in more details in § 2.4.

When more than one redshift per source is available, we almost entirely followed the choice of Xue et al. (2011). Briefly, the adopted redshift in this work is, in order of preference: secure (or upgraded to secure) spectroscopic redshift, photometric redshift from Luo et al. (2010), photometric redshift from Cardamone et al. (2010). XID 331 has no redshift information in the catalogues collected by Xue et al. (2011), but the inclusion of the other above-mentioned catalogues allows that source to be included in the sample. We assumed photometric redshift from Dahlen et al. (2010) for it. No spectroscopic redshift is reported in Xue et al. (2011) for XID 235, 458, 528 and 386; redshift for these sources were obtained by E. Vanzella (XID 235 and 458, private communication), Vanzella et al. (2010, XID 528) and Wuyts et al. (2009, XID 386). A high quality flag is associated to all these redshift in the original works; therefore, we consider them as "secure".

The redshift distribution of the sample is shown in Fig. 2.5. The mean and median redshifts are 4.0 and 3.7, respectively. Most of the sources with a spectroscopic redshift (13 out of 15) are at $z < 4$. We note the presence of 3 sources at $z > 5$, for which we investigated in more detail the redshift information. Two of them (XID 139 and 197) have only one photometric redshift ($z = 5.729$ and 6.071, respectively, from Luo et al. 2010, which also provides a secondary solution of $z = 4.385$ for XID 197), while XID 485 has two different photometric redshifts ($z = 7.62$ from Luo et al. 2010, with secondary solution $z = 3.309$, and $z = 4.42$ from Santini et al. 2009). These objects have no optical counterparts in the Great Observatories Origins Deep Survey (GOODS; XID 139 and 485) or Galaxy Evolution from Morphologies and SEDs (GEMS; XID 197, which is not in the GOODS field) observations, while they are significantly detected by the Infrared Array Camera (IRAC) in the near-IR. Thus, a high-redshift ($z > 3$) solution appears probable.

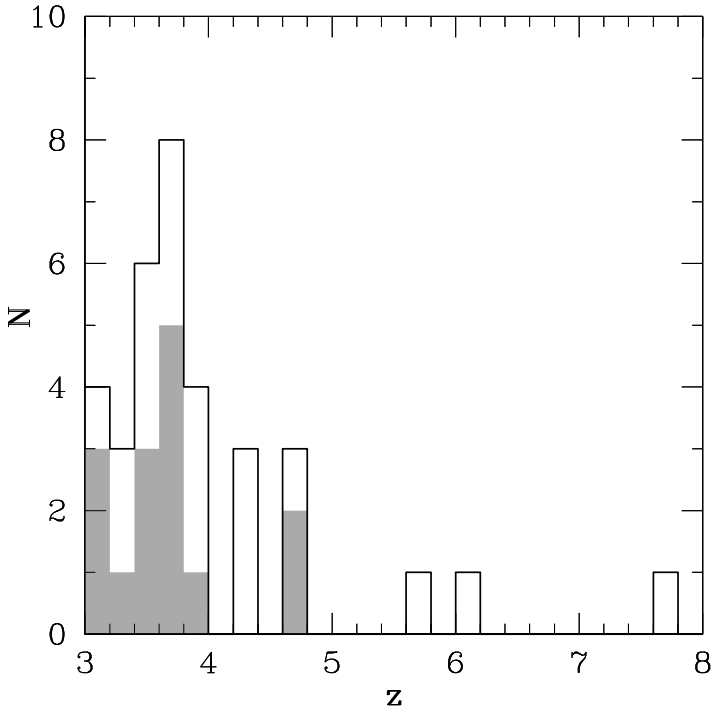

Figure 2.5: Redshift distribution of all 34 sources (solid line) and sources with spectroscopic redshift (shaded histogram) in the sample.

2.3 Data analysis

2.3.1 Spectral extraction procedure

Chandra data products for each of the 54 observations of the CDF-S are publicly available in the *Chandra* Data Archive[2]. We retrieved all data products (in particular, the evt2 files; i.e., the event list filtered on the good time intervals and status bits), but those related to the first two observations (for which the focal-plane temperature was $-110°$; Luo et al. 2008; Xue et al. 2011), and used them for spectral analysis. The sources were located in each of the 52 CDF-S images according to the coordinates provided by the main-source catalogue (Xue et al. 2011).

We selected the extraction regions for the source and background using SAOImage DS9[3] (Joye & Mandel 2003). Despite the *Chandra* PSF distortion with the off-axis angle, we always chose circular regions, provided that most of the source counts are included in the selected regions (probably this choice slightly lowers the signal-to-noise ratio). Small radii are required for faint sources in order to include a low number of background counts and therefore to increase the signal-to-noise ratio. Larger extraction radii can be used for bright sources. Also the position of the source in the field of view (f.o.v.) must be taken into account, because of the degradation of the PSF with the off-axis angle. For each source, we defined the same

[2]http://cxc.harvard.edu/cda/
[3]http://hea-www.harvard.edu/RD/ds9/

radius for the extraction region in all observations, considering the mean off-axis angle, provided by the main catalogue (Xue et al. 2011), and the relation between it and the encircled energy radius.

The off-axis angles for the sources in the sample are in the range of $1.42 - 10.1$ arcmin. We chose the extraction radius for each source to obtain an acceptable trade-off between the number of counts and the signal-to-noise ratio. As a general rule, for sources at off-axis angles $\lesssim 4$ arcmin we defined the extraction radius corresponding to an encircled energy fraction (EEF) of 90 per cent at 1.49 keV. For sources at a large off-axis angle ($\gtrsim 7$ arcmin), we considered radii in a range corresponding to 50-90 per cent EEF at 1.49 keV. For sources at intermediate off-axis angles (4-7 arcmin) we tried different extraction radii and chose those which provided the best trade-off between source and background counts. In a few cases we were forced to use smaller radii to avoid the contamination by sources close to the one we were considering. As a result, the extraction radii for the sample are in the range of $1.5 - 5$ arcsec. Regarding the background, we chose extraction regions as close as possible to the source, preferably on the same chip, in areas where no other source was detected. In particular, we excluded all the regions around detected sources with radii a factor ≥ 1.2 larger than the radii corresponding to 90% EEF at 6.4 keV (which are a factor > 1.5 larger than the radii corresponding to 90% EEF at 1.49 keV), in order to avoid contamination.

We used CIAO v4.3[4] to extract source and background spectra and to create the response files for each observation. In particular, we used the *specextract* tool which contains some improvements with respect to similar scripts in previous CIAO versions (e.g., it applies the energy-dependent aperture correction to the ARF file, accounting for the fraction of the PSF enclosed by a region). The spectra and response files of each source obtained for each observation were then added using tools provided in the FTOOLS package[5](i.e. *mathpha*, *addarf* and *addrmf*; Blackburn 1995), weighted by the exposure time of the individual observations.

The distribution of full-band ($0.5 - 7$ keV) net (i.e. background subtracted) counts (Fig. 2.6) is peaked near 100 counts (the mean and median values are 275 and 80, respectively). As expected, the bulk of the objects has a very low number of net counts: ~ 74 per cent of sources in the sample have less than 200 net counts. In order to maximise the signal-to-noise ratio (SNR), in most cases we fitted the spectra in a narrower energy range ($0.5 - 5$ keV), given that in the $5 - 7$ keV band the background is high. All the spectra were fitted using the Cash statistic (Cash 1979) to estimate the best-fitting parameters.

2.3.2 Spectral analysis

The spectral fitting was performed with XSPEC v12.5.1n[6] (Arnaud 1996). We assumed an absorbed power-law, parametrized by XSPEC models *powerlaw* and *zwabs*, as the starting spectral model. Galactic absorption was included assuming a column density in the range $(6 - 9) \times 10^{19}$ cm^{-2} (Kalberla et al. 2005).

The mean photon index is $\langle \Gamma \rangle = 1.51$ with dispersion $\sigma = 0.28$ for the 8 brightest

[4]http://cxc.harvard.edu/ciao/

[5]http://heasarc.nasa.gov/ftools/ftools_menu.html

[6]http://heasarc.nasa.gov/xanadu/xspec/

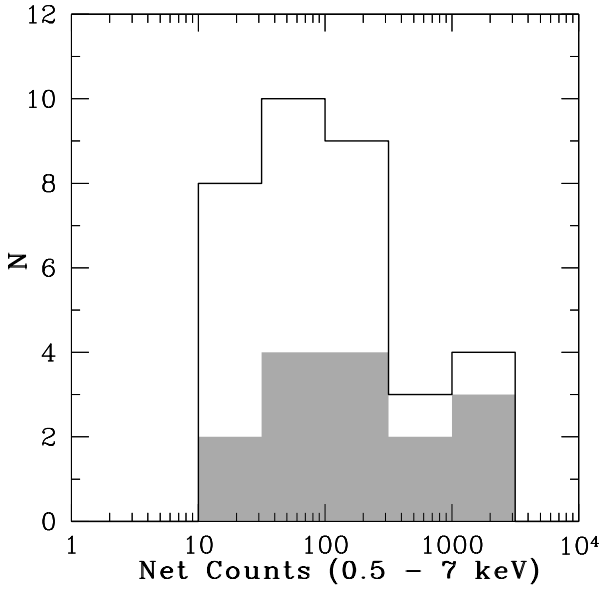

Figure 2.6: Full-band $(0.5-7$ keV) net counts distribution of all 34 sources (solid line) and sources with spectroscopic redshift (shaded histogram) in the sample.

sources (>250 net counts) and $\langle\Gamma\rangle = 1.63$ with $\sigma = 0.89$ for the whole sample. The average photon index is slightly flatter than what is typically measured in AGN samples ($\Gamma = 1.8$; e.g. Tozzi et al. 2006; Turner et al. 1997), but we emphasize that we are sampling the rest-frame $2-20$ keV band, and the likely presence of a reflection hump could flatten the observed X-ray spectrum. However, the measured value is statistically consistent with $\Gamma = 1.8$. Reasonable accuracy was obtained for the best-fitting parameters only for the 8 brightest sources (with a mean error on Γ and N_H of 14 and 45 per cent, respectively). The spectral parameters were not well constrained for almost all the other sources (error on Γ of 117 per cent; only upper limits on N_H for most objects) because of the poor spectral quality.

The difficulties in simultaneously constraining Γ and N_H for sources with few counts can be explained considering that these two spectral parameters have a significant degree of degeneracy. The problem is further complicated by the redshifted absorption cut-off, which moves outside the *Chandra* band-pass at $z \gtrsim 3$ even for significant degrees of obscuration, introducing further uncertainties.

To better constrain the column density, we fixed the photon index to $\Gamma = 1.8$ and repeated the spectral analysis on all the sources. This photon index is a widely used value for sources with low counting statistics, being considered the typical slope for AGN power-law emission in X-rays (e.g. Turner et al. 1997; Tozzi et al. 2006). The best-fitting parameters, fluxes and intrinsic luminosities derived using this spectral model are reported in Tab. 2.1. Similar results are obtained fitting simultaneously the source and background spectra. All the X-ray spectra of the sources, fitted using this model, are shown in Appendix A. The effects of assuming a flatter photon index ($\Gamma = 1.6$) will be discussed in § 2.3.3 and § 2.4.

Table 2.1: Main information and spectral parameters of the 34 high-redshift AGN.

XID	RA	DEC	z	r	C	N_H	$\frac{\text{CSTAT}}{\text{DOF}}$	SF	HF	HL	C
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)
27	52.96054	-27.87706	4.385	P8	5	68^{+41}_{-29}	71.7/113	2.45E-16	1.76E-15	2.80E44	77
100	53.01658	-27.74489	3.877	P8	3	38^{+27}_{-24}	69.0/92	1.48E-16	<7.55E-16	8.94E43	67
107	53.01975	-27.66267	3.808	P9	4	< 71	51.6/62	4.80E-16	<1.40E-15	1.52E44	36
132	53.03071	-27.82836	3.528	P8	3	13^{+16}_{-11}	69.0/65	1.13E-16	<3.39E-16	3.12E43	57
139	53.03496	-27.67975	5.729	P8	5	96^{+33}_{-35}	92.8/101	5.71E-16	3.64E-15	1.02E45	111
150	53.03996	-27.79847	3.337	P9	3	45^{+49}_{-39}	17.3/31	<2.24E-17	1.58E-16	1.40E43	26
170	53.04746	-27.87047	3.999	P8	4	35^{+4}_{-4}	308/359	1.63E-15	7.49E-15	9.43E44	1472
197	53.05771	-27.93344	6.071	P8	5	86^{+31}_{-30}	88.2/124	6.24E-16	3.27E-15	1.02E45	134
235	53.07029	-27.84564	3.712	S7	1	< 36	33.0/27	4.29E-17	7.95E-17	8.06E42	27
262	53.07854	-27.85992	3.66	S1	1	85^{+22}_{-20}	123.6/165	1.20E-16	1.53E-15	1.63E44	208
283	53.08467	-27.70811	3.204	P8	4	55^{+39}_{-22}	58.0/80	1.04E-16	9.54E-16	7.93E43	77
285	53.08558	-27.85822	4.253	P8	3	< 6	29.7/31	5.71E-17	<9.13E-17	1.24E43	18
331	53.10271	-27.86061	3.78	P10	4	< 23	48.1/36	4.61E-17	<7.39E-17	7.71E42	22
371	53.11158	-27.76789	3.101	P8	4	35^{+11}_{-10}	98.1/118	1.22E-16	7.70E-16	5.70E43	146
374	53.11204	-27.86072	3.724	P9	4	213^{+284}_{-114}	32.7/34	<3.59E-18	2.95E-16	4.14E43	22
386	53.11796	-27.73439	3.256	S6	1	< 7	62.6/64	8.85E-17	<1.42E-16	1.06E43	40
403	53.12196	-27.93883	4.762	S2	1	185^{+148}_{-76}	66.9/73	<5.46E-17	1.30E-15	2.68E44	36
412	53.12442	-27.85169	3.7	S1	1	82^{+12}_{-12}	208.0/225	2.08E-16	2.46E-15	2.85E44	373
458	53.13854	-27.82128	3.474	S7	1	93^{+74}_{-70}	14.7/25	1.59E-17	2.52E-16	2.64E43	27
485	53.14658	-27.87103	7.620	P8	4	406^{+192}_{-114}	74.2/94	5.75E-17	8.86E-16	5.00E44	87
521	53.1585	-27.73372	3.417	P8	4	< 23	47.6/50	8.91E-17	<2.40E-16	2.05E43	30
528	53.16158	-27.85606	3.951	S5	1	74^{+24}_{-24}	81.6/111	9.56E-17	8.93E-16	1.16E44	126
546	53.16533	-27.81419	3.064	S1	1	52^{+4}_{-4}	331.1/355	6.75E-16	6.26E-15	4.73E44	1072
556	53.17012	-27.92975	3.528	P8	3	97^{+9}_{-10}	285.6/293	7.59E-16	1.07E-14	1.14E45	718
563	53.17442	-27.86742	3.61	S1	1	6^{+2}_{-2}	257.6/307	2.07E-15	4.66E-15	4.45E44	1084
573	53.1785	-27.78411	3.193	S1	2	3^{+2}_{-2}	209.2/218	8.10E-16	1.65E-15	1.19E44	609
588	53.18467	-27.88103	3.471	S1	1	< 2	115.5/153	6.45E-16	1.03E-15	8.41E43	281
642	53.20821	-27.74994	3.769	P9	3	< 13	55.7/65	9.17E-17	<1.48E-16	1.54E43	18
645	53.20933	-27.88119	3.47	S3	2	15^{+2}_{-2}	371.4/384	3.16E-15	1.03E-14	9.24E44	2053
651	53.21529	-27.87033	4.658	P8	3	151^{+39}_{-35}	115.8/132	1.33E-16	2.38E-15	4.64E44	146
674	53.24004	-27.76361	3.082	S3	1	< 7	120.3/146	6.95E-16	1.39E-15	8.56E43	199
700	53.2625	-27.86308	4.253	P8	5	18^{+16}_{-14}	73.8/90	7.43E-16	2.23E-15	3.11E44	102
717	53.28	-27.79892	4.635	S4	2	87^{+65}_{-51}	47.2/73	1.21E-16	1.02E-15	1.85E44	44
730	53.29587	-27.79317	3.724	S4	1	60^{+49}_{-39}	59.6/83	<1.81E-16	<1.46E-15	1.65E44	37

(1) Identification number from Xue et al. (2011); (2) Right ascension and (3) declination (J2000) from Xue et al. (2011); (4) adopted redshift (see § 2.2); (5) reference for the adopted redshift - S1, S2, S3, S4, S5, S6 and S7: spectroscopic redshift from Szokoly et al. (2004), Vanzella et al. (2008), Popesso et al. (2009), Silverman et al. (2010), Vanzella et al. (2010), Wuyts et al. (2009) and E. Vanzella (private communication), respectively; P8, P9 and P10: photometric redshift from Luo et al. (2010), Cardamone et al. (2010) and Dahlen et al. (2010, weighted solution), respectively; (6) criterion satisfied for source inclusion in the high-redshift sample (see § 2.2); (7) best-fitting N_H as derived in § 2.3.2, in units of 10^{22} cm^{-2}. We define N_H to be constrained if the lower limit (at the 90 per cent c.l.) on the best-fitting value is larger than zero, otherwise we report its upper limit. (8) best-fit value of Cash statistic over degrees of freedom; (9) soft-band ($0.5-2$ keV) and (10) hard-band ($2-10$ keV) flux or upper limits from the best-fitting spectral model, in units of erg cm^{-2}s^{-1}; (11) intrinsic (i.e. absorption-corrected) rest-frame $2-10$ keV luminosity, in units of erg s^{-1} (typical errors range from ~ 10 per cent for the brightest sources to a factor of ~ 2 for the faintest ones, fixing the redshift to the value reported in column 4); (12) full-band ($0.5-7$ keV) net (i.e. background-subtracted) counts, as obtained from the spectral extraction. The full-band net-count number is only representative, being the energy range considered during the spectral fitting narrower, in most cases, in order to maximise signal-to-noise ratio (SNR).

Comparison with *XMM* results

A ultra-deep *XMM-Newton* survey (Comastri et al. 2011; Ranalli et al. 2013) was performed over an area encompassing the CDFS field, for a total of ~ 3 Ms nominal exposure. Fourteen *XMM* sources are associated with *Chandra*-detected high-redshift AGN in the main catalogue (Ranalli et al. 2013). We could therefore exploit the high throughput of *XMM-Newton* to investigate in more details the spectral parameters of these objects. In order to avoid strong background-related issues, we limited ourself to the 7 brightest AGN by considering only sources with flux $F_{2-10\ \mathrm{keV}} > 2 \times 10^{-15}\,\mathrm{erg\,cm^{-2}s^{-1}}$ in the X-ray point-source catalogue. Spectra were binned to have at least 50 counts per energy bin[7] and the χ^2 statistics was assumed during the fit. We performed a simultaneous fit to the *pn* and MOSs X-ray spectra using the same spectral models described in § 2.3.2; i.e. an absorbed power-law, first with Γ free to vary (Tab. 2.2) and then fixed to $\Gamma = 1.8$ (Tab. 2.3). In Fig. 2.7 and 2.8 we compare the best-fitting column density and hard-band luminosity derived from *XMM* and *Chandra* spectra in the two cases, respectively. The *XMM* results are generally consistent with the *Chandra* values, especially assuming the same, fixed photon index. As *Chandra* and *XMM* did not observed the field simultaneously, we checked for the possible variability of the spectra by comparing the soft-band fluxes derived from the spectral analysis in the two cases and we found them to be consistent within the errors.

We also repeated the procedure without subtracting the background spectra from the source spectra, but performing a simultaneous background and source spectra fitting. This procedure is appropriate when dealing with background-dominated spectra. First, we fitted a phenomenological model, which included cosmic and instrumental features, extracted from a region close to the source. Then this model, with all the parameter fixed but the normalization, was added during the fit of the source spectra, from which the background spectra were not subtracted. The results are broadly consistent with those reported in Tab. 2.2 and Tab. 2.3. In conclusion, despite the larger effective area of *XMM*-Newton with respect to *Chandra*, its higher level of background prevented us from deriving the value of the spectral parameters more precisely than using *Chandra* data. No significant features were observed in the *XMM* spectra of sources in our sample.

2.3.3 Column density distribution

The distribution of the best-fitting N_H (Fig. 2.9) derived from spectral analysis, assuming an absorbed power-law with fixed $\Gamma = 1.8$ as fitting spectral model, strongly suggests the presence of an important fraction ($\approx 68\%$) of highly obscured sources ($N_H > 10^{23}$ cm^{-2}). This result could be affected by a systematic N_H overestimate due to the high redshift and low counting statistics, characteristic of the sample: the photoelectric cut-off in spectra of mildly-obscured ($10^{22} < N_H < 10^{23}$ cm^{-2}) sources is redshifted out of the *Chandra* band-pass already at $z = 3-4$, making very difficult the search for obscuration. Indeed, this type of spectrum might be indistinguishable from a spectrum of an unabsorbed source at high redshift. Furthermore,

[7]The same results, within the errors, were derived by binning the *XMM* spectra so that the signal-to-noise ratio in each energy bin was ≥ 3.

Table 2.2: Spectral parameters of the 7 brightest high-redshift *XMM* sources with a *Chandra* counterpart selected in the CDFS. Photon index was left free to vary.

PID	XID	z	Γ	$N_H[10^{22}cm^{-2}]$	χ^2/DOF	SF	HF	HL
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
64	556	3.528	$1.07^{+0.45}_{-0.38}$	41^{+28}_{-20}	97/101	5.64e-16	8.14e-15	2.60e44
107	645	3.47	$1.66^{+0.15}_{-0.14}$	8^{+2}_{-2}	105/85	2.12e-15	6.37e-15	4.57e44
116	170	3.999	$1.64^{+0.19}_{-0.17}$	24^{+7}_{-6}	90/83	1.35e-15	6.16e-15	5.84e44
120	563	3.61	$1.67^{+0.16}_{-0.13}$	< 2	76/83	1.64e-15	3.35e-15	2.61e44
144	412	3.7	$2.55^{+1.80}_{-1.31}$	88^{+93}_{-65}	111/113	2.21e-16	1.24e-15	5.00e44
180	546	3.064	$1.62^{+0.37}_{-0.33}$	53^{+18}_{-15}	50/67	5.66e-16	6.54e-15	3.81e44
210	573	3.193	$2.04^{+0.46}_{-0.38}$	6^{+5}_{-4}	114/11	7.68e-16	1.65e-15	1.35e44

(1) *XMM* identification number from Ranalli et al. (2013); (2) *Chandra* identification number from Xue et al. (2011); (3) redshift; (4) best-fitting photon index; (5) best-fitting column density in units of 10^{22} cm^{-2}; (6) best-fit value of χ^2 over degrees of freedom; (7) soft-band ($0.5-2$ keV) and (8) hard-band ($2-10$ keV) flux from the best-fitting spectral model, in units of erg cm^{-2}s^{-1}; (9) intrinsic (i.e. absorption-corrected) rest-frame $2-10$ keV luminosity, in units of erg s^{-1}.

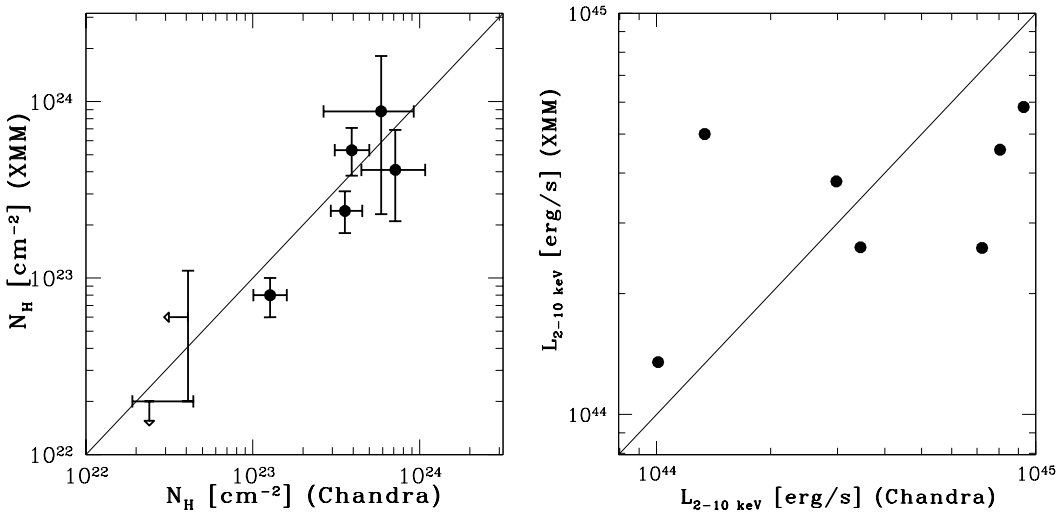

Figure 2.7: Best-fitting column density (left panel) and hard-band luminosity (right panel) for the 7 brightest high-redshift sources detected by both *XMM* and *Chandra* in the CDFS. The photon index was left free to vary. Straight line is the 1:1 relation.

Table 2.3: Spectral parameters of the 7 brightest high-redshift *XMM* sources with a *Chandra* counterpart selected in the CDFS. Photon index was fixed to $\Gamma = 1.8$.

PID	XID	z	$N_H[10^{22}cm^{-2}]$	χ^2/DOF	SF	HF	HL
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
64	556	3.528	81^{+16}_{-14}	103/102	5.36e-16	6.34e-15	6.60e44
107	645	3.47	10^{+1}_{-1}	107/86	2.13e-15	5.76e-15	5.09e44
116	170	3.999	30^{+4}_{-4}	92/84	1.35e-15	5.67e-15	7.07e44
120	563	3.61	2^{+1}_{-1}	78/84	1.65e-15	2.97e-15	2.80e44
144	412	3.7	52^{+23}_{-19}	111/114	2.25e-16	1.59e-15	1.75e44
180	546	3.064	60^{+8}_{-7}	51/68	5.66e-16	6.19e-15	4.78e44
210	573	3.193	4^{+2}_{-2}	115/112	7.65e-16	1.62e-15	1.17e44

(1) *XMM* identification number from Ranalli et al. (2013); (2) *Chandra* identification number from Xue et al. (2011); (3) redshift; (4) best-fitting column density in units of 10^{22} cm^{-2}; (5) best-fit value of χ^2 over degrees of freedom; (6) soft-band ($0.5-2$ keV) and (7) hard-band ($2-10$ keV) flux from the best-fitting spectral model, in units of erg cm^{-2}s^{-1}; (8) intrinsic (i.e. absorption-corrected) rest-frame $2-10$ keV luminosity, in units of erg s^{-1}.

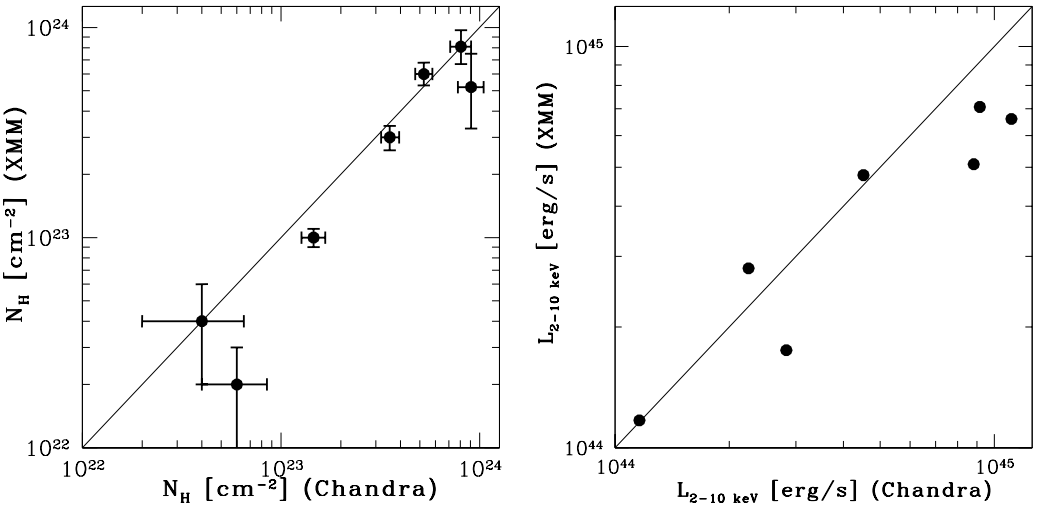

Figure 2.8: Same as Fig. 2.7, but the photon index was fixed to $\Gamma = 1.8$.

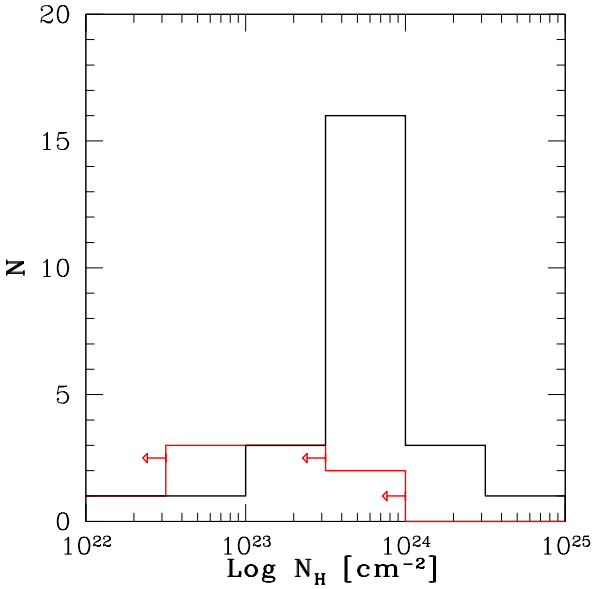

Figure 2.9: Observed distribution of the best-fitting N_H. The black line represents the constrained values; the red line with leftward pointing arrows represents the upper limits.

statistical fluctuations, enhanced by the low fluxes, are expected in the spectra and could simulate a spurious photoelectric cut-off at high rest-frame energies (Tozzi et al. 2006).

In order to estimate the probability (P_{ij}) that the X-ray spectrum of a source with intrinsic N_H in a given range j is fitted with a best-fitting column density constrained in a range i, we made extensive use of simulations. As in Tab. 2.1, we define N_H to be constrained if the 90% c.l. lower limit on the best-fitting value is larger than zero. We divided the sample in 5 N_H bins, as reported in Tab. 2.4. The number of sources N_i observed in each absorption bin i is then given by:

$$N_i = \sum_j x_j P_{ij} \tag{2.1}$$

where x_j is the number of sources with intrinsic N_H in bin j and N_i is the number of sources observed with best-fitting N_H in bin i, for i, j = A, B, C, D, E. In this framework, N_i is known (black histogram in Fig. 2.9) and P_{ij} can be derived from simulations.

We simulated 1000 spectra with $\Gamma = 1.8$ and 100 net counts (close to the median value of our sample), intrinsically obscured by a column density in the j-th N_H bin (for each bin we used the value reported in the third column of Tab. 2.4), considering sources at $z = 4$ (similar to the characteristic redshift of our sample), and simply counted how many times the simulated spectra are fitted with a best-fitting N_H constrained in the i-th bin (see Appendix B for the detailed description of the simulation procedure). Bins A, B and C are indistinguishable at $z > 3$ and hence are merged into a single bin (representing unobscured or mildly-obscured sources with $N_H < 10^{23}$ cm^{-2}), hereafter referred to as (ABC).

Table 2.4: N_H bins used in simulations

i,j	range: $\log(\frac{N_H}{\mathrm{cm}^{-2}})$	sim: $\log(\frac{N_H}{\mathrm{cm}^{-2}})$
A	< 21	0.0
B	21-22	21.5
C	22-23	22.5
D	23-24	23.5
E	>24	24.5

Sources in the i-th or j-th N_H bin (first column), which accounts for the N_H range shown in the second column, are simulated assuming the column density reported in the third column.

Table 2.5: Final probability factors P_{ij} that a source with intrinsic N_H in a given bin j is observed in a bin i. These are derived from simulations using $\Gamma = 1.8$ and $\Gamma = 1.6$.

P_{ij}	j=(ABC)	j=D	j=E
	$\Gamma = 1.8$		
i=(ABC)	0.361	0.0	0.0
i=D	0.639	1.0	0.0
i=E	0.0	0.0	1.0
	$\Gamma = 1.6$		
i=(ABC)	0.147	0.0	0.0
i=D	0.853	1.0	0.0
i=E	0.0	0.0	1.0

The value of each P_{ij} is shown in Tab. 2.5. We immediately note that no spurious effect seems to be important in bins D and E (i.e. spectra of intrinsically strongly obscured sources are efficiently fitted by a model obscured by a column density within the correct N_H bin) while, as expected, the best-fitting N_H is overestimated in most cases ($\approx 64\%$) when the intrinsic value is in the bin (ABC).

By inverting eq. (2.1), the intrinsic N_H distribution can now be derived (Fig. 2.10): $x_{\,(\mathrm{ABC})} \approx 5.5$, $x_{\,\mathrm{D}} \approx 15.5$ and $x_{\,\mathrm{E}} = 4$, to be compared with the observed distribution: $N_{\,(\mathrm{ABC})} = 2$, $N_{\,\mathrm{D}} = 19$ and $N_{\,\mathrm{E}} = 4$. If we conservatively count all the 9 sources for which only upper limits could be obtained (red line in Fig. 2.9) in the first bin, the resulting N_H distribution is the one shown in Fig. 2.11. We added the predictions of the Gilli, Comastri, & Hasinger (2007) X-ray background

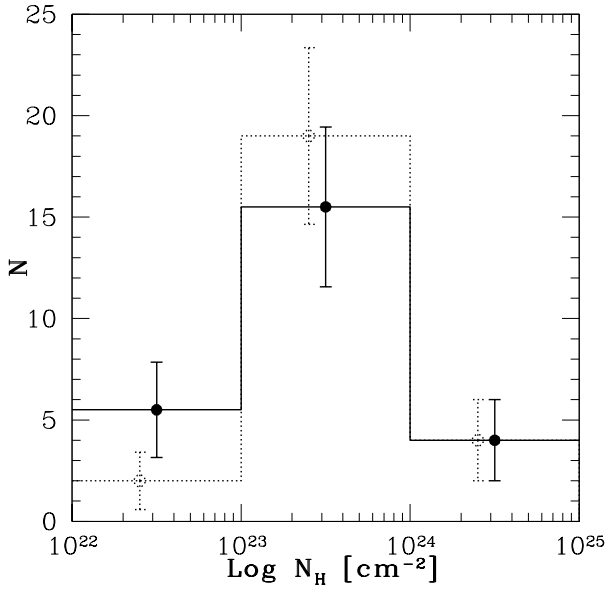

Figure 2.10: Intrinsic N_H distribution (solid line) with statistical errors, derived correcting the observed one (dotted line, plotted in the same bins, with statistical errors slightly shifted for visual purpose), resulting from spectral analysis assuming $\Gamma = 1.8$, with the probability factors obtained from simulations (see § 2.3.3 and Appendix B). Only constrained values of N_H are considered.

(XRB) synthesis model[8] computed with the POMPA[9] tool for each N_H bin, considering sources with luminosity $42 \leq \mathrm{Log}\frac{L_{0.5-2\ \mathrm{keV}}}{\mathrm{erg\ s^{-1}}} \leq 47$ in a redshift range between $z_{min} = 3$ and $z_{max} = 8$. A high-redshift exponential decline in the space density of AGN, similar to that of bright quasars (Gilli et al. 2011), is also considered. The assumption of the decline affects only the normalization and not the shape of the column density distribution. The predictions are corrected for the sky coverage (i.e. the sky area in physical units covered by the survey at different fluxes, see next section), to take into account the smaller sky area covered by the survey at low fluxes. Obscured sources have a lower flux than unobscured or less obscured ones with the same luminosity at the same redshift and, hence, can be detected over a smaller, deeper area.

The N_H distribution, corrected for the spurious overestimate of the column density, shows a number of strongly obscured sources that is slightly larger than the predictions of the model we considered. Conversely, the number of unobscured or mildly-obscured sources is smaller than the predicted number. However, we find an

[8]The Gilli et al. (2007) model was developed to reproduce the XRB, by estimating the number of obscured Compton-Thin AGN from the comparison of soft (Hasinger et al. 2005) and hard (Ueda et al. 2003; La Franca et al. 2005) X-ray luminosity functions. Their absorption distribution was found by fitting the number counts in different energy bands. Finally, the number of Compton-Thick AGN was computed by adding as many source as as needed to fit the residual XRB peak. The ratio between obscured Compton-Thin and unobscured AGN was assumed to be constant with redshift and to decrease from ~ 4 at low luminosities to ~ 1 at high luminosities.

[9]POrtable Multi-Purpose Application for AGN counts
(http://www.bo.astro.it/∼gilli/counts.html)

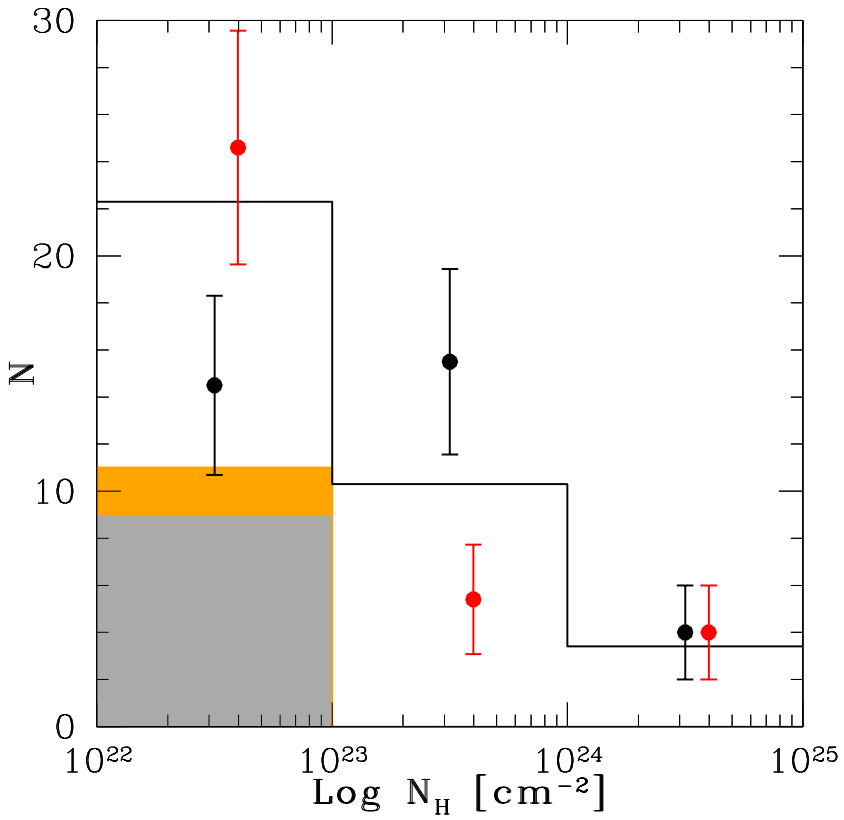

Figure 2.11: Intrinsic N_H distribution assuming $\Gamma = 1.8$ (black points) and $\Gamma = 1.6$ (red points, which are slightly shifted from the centre of the bins for visual purposes). The first bin includes sources with $N_H < 10^{23}$ cm^{-2} (i.e. consistent with no or "moderate" absorption) and all the sources for which only upper limits on N_H were obtained (9 and 11 sources in the two cases, represented as grey and orange shaded areas, respectively). The third bin contains sources with $N_H > 10^{24}$ cm^{-2}. Poissonian errors are computed for the number of sources in each bin. The histogram is the predictions computed with the POMPA tool and corrected for the sky coverage of the survey (see text). The predictions are not normalized to the observed number of sources.

agreement between the observations and the predictions at the 2σ confidence level (the χ^2 test returns a value of 6.0 for 3 degrees of freedom and no free parameter).

Following the results of § 2.3.2 where, considering an absorbed power-law with Γ free to vary as starting spectral model, we derived a mean photon index $< \Gamma > \approx 1.6$, we investigated the effects of assuming a flatter photon index on the column density distribution. Therefore, we fitted again the X-ray spectra of all sources in the sample using the same spectral model with photon index fixed to $\Gamma = 1.6$. The resulting N_H distribution was corrected for spurious overestimates using the probability factors (Tab. 2.5) obtained from simulations, similarly to the $\Gamma = 1.8$ case, but now using $\Gamma = 1.6$ to simulate and fit the spectra. Red points in Fig. 2.11 represent the intrinsic N_H distribution in the $\Gamma = 1.6$ case. We derived upper limits on N_H for two sources (XID 132 and 700, which were previously counted in bin D), while their best-fitting column density was constrained using $\Gamma = 1.8$. This result produces a different number of sources in the two shaded histograms in Fig. 2.11. Therefore, the number of objects in each bin of the corrected N_H distribution is, in this case, $x_{\,(\mathrm{ABC})} \approx 13.6$, $x_{\,\mathrm{D}} \approx 5.4$ and $x_{\,\mathrm{E}} = 4$. If we place all the 11 sources with an upper limit on N_H in the first bin, as in Fig. 2.11, the distribution is fully consistent within 2σ with the prediction of the model ($\chi^2 = 4.75$ with 3 degrees of freedom and no free parameter). We underline that the model predictions are computed with the same level of incompleteness as the observed sample.

2.3.4 Number counts

We derived the logN-logS of the high-redshift sample in the soft band, where the survey sensitivity is the highest, folding the number N of sources with flux larger than S with the survey sky-coverage at that flux. We assumed the soft fluxes reported in Tab. 2.1 derived from our spectral analysis. The sky-coverage was obtained by multiplying the nominal sky area covered by the survey (≈ 0.13 deg^2; Xue et al. 2011) and the fraction of the f.o.v. covered at different fluxes, derived applying the CIAO tool DMIMGHIST on the sensitivity map (Xue et al. 2011), which assigns to each point of the f.o.v. the flux limit of the survey at that position. In order to be consistent with the computation of the sky coverage (derived in the $0.5 - 2$ keV band), we excluded four ($\approx 12\%$ of the total sample) sources which were not detected in the soft band (see Tab. 2.1).

In Fig. 2.12 we compare our data with those provided by Kalfountzou et al. (2014) for *Chandra*-COSMOS and ChAMP surveys, Fiore et al. (2012) and Lehmer et al. (2012) for the 4 Ms CDF-S and with the XMM (Brusa et al. 2009) and *Chandra* (Civano et al. 2011) COSMOS data. The predictions of the Gilli et al. (2007) XRB synthesis model (solid and dashed lines) and the Aird et al. (2010) X-ray luminosity function (XLF) evolutionary model (dotted line) are also shown. The Gilli et al. (2007) model is based on a luminosity dependent density evolution (LDDE) model of the XLF, and it allows the presence of a high-redshift ($z > 2.7$) decline in the AGN space density to be taken into account. The Aird et al. (2010) model assumes a luminosity and density evolution (LADE) of the XLF.

Our results for the $z > 3$ sample (*left panel*) are in good agreement with the Gilli et al. (2007) model predictions assuming an exponential decline in high-redshift AGN space density (continuous line), down to flux $F_{0.5-2\ \mathrm{keV}} \approx 4 \times$

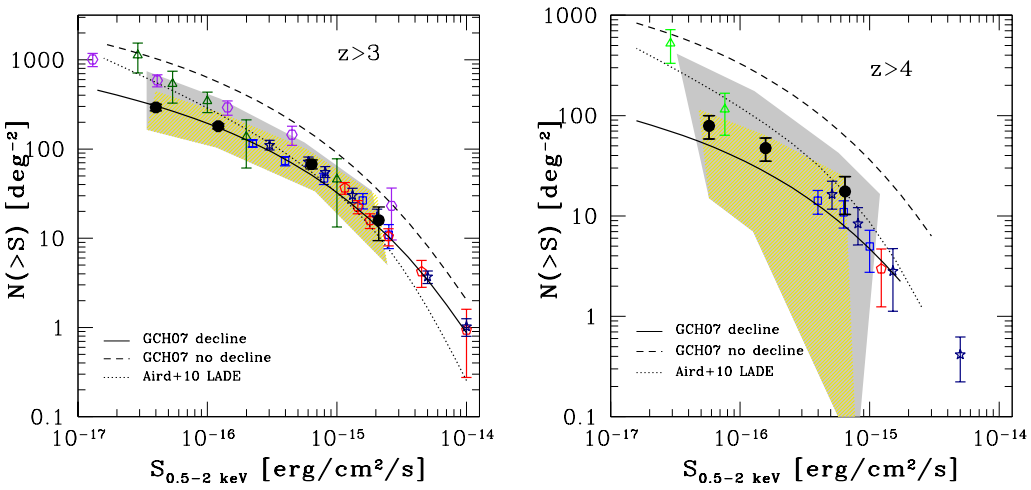

Figure 2.12: Binned LogN-logS of the high-redshift sample, considering all sources ($z > 3$, *left panel*) and only those at $z > 4$ (*right panel*), with associated Poissonian errors. Black filled circles are from this work, dark-blue stars are from Kalfountzou et al. (2014), green triangles are from Fiore et al. (2012), blue squares are from Civano et al. (2011), red pentagons are from Brusa et al. (2009) and purple hexagons are from Lehmer et al. (2012). Empty black circles in the right panel represent our data if the three sources at the highest redshifts were placed at $3 < z < 4$ (see text). The predictions of the Gilli et al. (2007) XRB synthesis model, computed with the POMPA tool, either assuming (solid line) or not (dashed line) an exponential decline in the AGN space density at high redshift, and Aird et al. (2010) model are also shown. The yellow shaded area represents the regions in which the LogN-LogS would be placed if more conservative or shallower selection criteria were used as well as the grey area, which also includes the 31 soft-band detected sources with no redshift in the considered catalogues (see § 2.4).

10^{-17} erg cm^{-2}s^{-1}, a factor ≈ 10 fainter with respect to previous surveys (i.e. down to $L_X \approx 10^{43}$ erg s^{-1}, the minimum luminosity detectable at $z > 3$ in the 4 Ms CDF-S): we extend to lower fluxes the results obtained in the COSMOS field by Brusa et al. (2009) and Civano et al. (2011), confirming an evolution at high redshift similar to that of the more luminous AGN (Gilli et al. 2011).

Considering only the $z > 4$ subsample (Fig. 2.12, *right panel*, filled circles), our data lie above the model predictions. We include here the presence of 3 sources at $5 < z \lesssim 7.7$, whose redshifts are determined on the basis of relatively uncertain photometric information. If these sources were placed at $3 < z < 4$, a good agreement would be obtained (Fig. 2.12, *right panel*, open circles).

The scenario without the decline (Fig. 2.12, dashed lines) is excluded by our data both at $z > 3$ and $z > 4$.

2.4 Discussion

The N_H distribution derived from spectral analysis, corrected for the overestimate of the column density, suggests the presence of a significant fraction ($\sim$ 57 per cent) of strongly obscured ($N_H > 10^{23}$ cm^{-2}) AGN at $z > 3$. This result was obtained by fitting the X-ray spectra with an absorbed power-law with $\Gamma = 1.8$, i.e. the typical photon index observed in large AGN samples (e.g. Turner et al. 1997; Tozzi et al. 2006). Adopting $\Gamma = 1.6$, we derived a slightly different intrinsic N_H distribution, with more unobscured sources with respect to the previous case. Since the current data are sampling a rest-frame energy range up to at least $E_{int} \approx 20$ keV, a hardening of the X-ray spectra due to reflection might be expected at observed energies $E_{obs} \approx 4 - 5$ keV. However, the low photon-counting statistics prevented us from using relatively complex spectral models and the $\Gamma = 1.6$ index may simply incorporate the hardening due to reflection. Therefore, the "true" N_H distribution is expected to lie between the two distributions derived assuming $\Gamma = 1.8$ and $\Gamma = 1.6$ (Fig. 2.11). To further investigate the possible impact of a reflection component, we computed the stacked, rest frame, X-ray spectra (using the method described in Iwasawa et al. 2012) dividing the sample in four N_H bins ($\frac{N_H}{10^{23}\ \mathrm{cm^{-2}}} < 1.3$; $1.3 < \frac{N_H}{10^{23}\ \mathrm{cm^{-2}}} < 6$; $6 < \frac{N_H}{10^{23}\ \mathrm{cm^{-2}}} < 9$; $\frac{N_H}{10^{23}\ \mathrm{cm^{-2}}} > 9$; see table 1). The resulting best-fit slope, obtained by a joint fit of the the various N_H bins and representing the intrinsic continuum above 10 keV, is $\Gamma \simeq 1.69^{+0.07}_{-0.07}$. This slope is suggestive of a limited contribution from reflection and is well within the values adopted to compute the absorption distribution (Fig. 8). Independently from the assumed spectrum, the data are in agreement at the 2σ confidence level with the prediction of the Gilli et al. (2007) XRB synthesis model, assuming an exponential decline in the AGN space density at $z > 2.7$ (Fig. 2.11). The best-fitting column densities of the 19 high-luminosity ($L_{2-10\ \mathrm{keV}} > 10^{44}$ erg s^{-1}) sources are well constrained at high values ($\sim$ 84 per cent of them have nominal $N_H > 10^{23}$ cm^{-2}). Applying the correction factors derived assuming $\Gamma = 1.8$ (see § 2.3.3), we derived that the intrinsic fraction of obscured sources ($N_H > 10^{23}$ cm^{-2}) in this subsample is $\sim$ 66 per cent. These results agree with the recently-proposed scenario in which the evolutionary paths of low and high-luminosity AGN are different (see § 1.4.1), causing a different evolution (more evident for luminous sources) of the obscured AGN fraction with redshift.

Fig. 2.12 shows the number counts derived in this work compared to previous determinations from the literature and theoretical models. The CDF-S source counts down to $F_{0.5-2\ \mathrm{keV}} \approx 4\times10^{-17}$ erg cm^{-2}s^{-1} lie on the XRB model if an exponential decline of the AGN X-ray luminosity function is assumed, similarly to *XMM* (Brusa et al. 2009) and *Chandra* (Civano et al. 2011) COSMOS results at brighter fluxes. In particular, we note the agreement at bright fluxes with the results from COSMOS surveys, for which a major effort was made to obtain spectroscopic and photometric redshifts (Salvato et al. 2009; Brusa et al. 2009; Salvato et al. 2011, and references therein). Our results are consistent also with the Aird et al. (2010) model if the least conservative bounds (grey area in Fig. 2.12) are assumed (see below for the computation of the error budget). At fluxes $2\times10^{-16} \leq F_{0.5-2\ \mathrm{keV}} \leq 2\times10^{-15}$ erg cm^{-2}s^{-1} (i.e. for the brightest sources in the high-redshift sample) the prediction of this model is very similar to the Gilli et al. (2007) model, but at fainter fluxes the two models deviate significantly. Our results better trace the Gilli et al. (2007) model, while data from Fiore et al. (2012) and Lehmer et al. (2012) are placed on the Aird et al. (2010) model.

There is consistency in the number counts that we present with data from Fiore et al. (2012) and Lehmer et al. (2012) down to $F_{0.5-2\ \mathrm{keV}} \approx 10^{-16}$ erg cm^{-2}s^{-1}, while our results are a factor of $\sim 2-3$ lower at fainter fluxes. This is not surprising considering that different selection techniques are applied, ours being more conservative than the ones used in those works. Indeed, the logN-logS in Fiore et al. (2012) is derived for 54 optically selected sources at $z > 3$, using spectroscopic and photometric redshifts and drop-out techniques and applying a new detection algorithm on the 4 Ms CDF-S observations, which made possible to reach a flux limit of $F_{0.5-2\ \mathrm{keV}} \approx 1.2\times10^{-17}$ erg cm^{-2}s^{-1}. The samples considered in Lehmer et al. (2012) and in this work are collected from the same catalogue of X-ray selected sources. However, Lehmer et al. (2012) assumed the redshift adopted in Xue et al. (2011), while we collected additional information from other catalogues and used a stricter selection procedure (as described in § 2.2). Therefore, we rejected a number of sources which are included in Lehmer et al. (2012), and we were able to collect redshift information for some objects which have no redshift in Xue et al. (2011), as in the case of XID 331. We would obtain results largely consistent with Lehmer et al. (2012) using their sample.

To compute the "error budget" on the number counts, we investigate how many sources would be included in the sample if we relax the third, fourth and fifth criteria described in § 2.2. Relaxing the third criterion and including in the final sample soft X-ray sources for which $N_{phot,>3} - N_{phot,<3} \geq 0$, the net gain in the final sample would be of 4 sources. As for the fourth criterion, 14 sources were excluded from the sample because only one of the two available photometric redshift is at $z > 3$. We consider these objects to have a 50% chance of being at $z > 3$, and randomly include seven of them in the error budget. Finally, if we relax the fifth criterion and include in the sample all sources with only one photometric redshift, with the condition of being at $z > 3$, regardless of its error, we would add 4 sources to the final sample. If all these changes were adopted, the sample would consist of 49 sources.

The gold shaded area in Fig. 2.12 indicates the regions in which the logN-logS

would lie considering the error budget [10]. For the $z > 3$ number counts (upper panel), the lower bound is obtained if only sources selected via the first, second and third criteria of § 2.2 are included in the high-redshift sample (i.e. sources with spectroscopic redshift or $N_{phot,>3} - N_{phot,<3} \geq 3$). This is clearly a very conservative assumption. The upper bound is obtained if we relax the third, fourth and fifth criteria, as discussed above, adding 15 sources in total.

Four out of these 15 sources would be at $z > 4$: two would be selected relaxing the fifth criterion and two relaxing the fourth criterion. As previously described, we randomly assumed half of the latter sources to be at $z > 4$. The resulting 3 sources were added to the $z > 4$ sample to compute the upper-bound of the gold shaded area (bottom panel in Fig. 2.12), which accounts for the error budget of the $z > 4$ number counts. The lower bound is obtained, as before, considering only the first, second and third criteria, and consists of one source with a spectroscopic redshift $z > 4$ (XID 717, while we discarded XID 403 because it is not detected in the soft band) and 2 sources selected through the third criterion. As expected, because of the low number of sources detectable at $z > 4$, only loose constraints can be placed on the LogN-LogS. Larger samples of X-ray selected AGN at $z > 4$ would be needed to draw more solid conclusions.

Thirty-nine sources have no redshift in any catalogue that we considered. Thirty-one of them are detected in the soft band in Xue et al. (2011). A not negligible fraction of them are probably AGN at $z > 3$. The grey shaded regions in Fig. 2.12 show how the gold area would enlarge if all these sources would be accounted for. Therefore, these areas represent the hypothetical and unlikely cases in which all the 31 sources were at $z > 3$ ($z > 4$).

2.5 Summary and conclusions

The detection of a significant number of $z > 3$ AGN down to $L_X \approx 10^{43}$ erg s^{-1} is made possible by the faint flux limit reached in the 4 Ms CDF-S observations (Xue et al. 2011).The large multi-wavelength coverage of the CDF-S allows us to retrieve spectroscopic and photometric redshifts for most of the sources in the field. In this framework, physical and evolutionary properties of a relatively large sample of high-redshift AGN can be studied.

The sample consists of 34 sources in the range $3 < z \leq 7.6$, with a median redshift of $z = 3.7$. About 45 per cent of them (15 sources) have a spectroscopic redshift.

X-ray spectra of all the sources were extracted from the 4 Ms CDF-S observations and fitted with an absorbed power-law spectral model with photon index fixed to $\Gamma = 1.8$. The low photon-counting statistics prevents us from assuming a more complex spectral model: the median full-band ($0.5-7$ keV) net-count number is 80. The median rest-frame absorption-corrected luminosity of the sample is $L_{2-10 \text{ keV}} \approx 1.5 \times 10^{44}$ erg s^{-1}.

We summarise the conclusions reached in this chapter as follows:

1. The observed column density distribution is strongly peaked between

[10] We assigned the soft X-ray fluxes reported in Xue et al. (2011) to the sources not included in the high-redshift sample for the computation of all shaded areas.

$23 < \log N_H < 24$. The low number of counts and the high redshift of the source sample may lead to overestimate the intrinsic absorption. To account for this effect, we ran extensive X-ray spectral simulations and derived correction factors to apply to the observed distribution. Therefore we obtained an estimate of the intrinsic N_H distribution. The results are consistent with a non-evolving obscured fraction of AGN with $L_x \approx 10^{44}$ erg s^{-1} with redshift within 2σ.

2. The number counts of our $z > 3$ AGN sample are consistent with a decline in the AGN space density at high redshift, extending at faint soft fluxes ($F_{0.5-2\ \mathrm{keV}} \approx 4 \times 10^{-17}$ erg cm^{-2}s^{-1}) the behaviour determined at brighter fluxes found by Brusa et al. (2009) and Civano et al. (2011) in the COSMOS field. We used a more conservative approach with respect to Fiore et al. (2012) and Lehmer et al. (2012): if we significantly relax our selection criteria, our results are in a relatively good agreement at bright fluxes with their data, while at the faint end the discrepancy is a factor ~ 2.

   These results would be improved by increasing the size and the reliability of the sample (e.g. gathering additional reliable redshifts via ultradeep optical-IR spectroscopic campaigns) and the quality of the X-ray spectra (e.g. increasing the *Chandra* exposure in this field). The accuracy of the photometric redshifts would be significantly improved by adding the deep Cosmic Assembly Near-IR Deep Legacy Survey (CANDELS) H and J-band data (Grogin et al. 2011; Koekemoer et al. 2011).

Chapter 3

The Hard X-ray Luminosity Function of High-Redshift $3 < z \lesssim 5$ AGN

In this chapter, we assembled a sample of X-ray selected $z > 3$ AGN and we used it to derive the hard-band ($2-10$ keV) X-ray luminosity function (HXLF) of high-redshift AGN. Objects collected from the *Chandra*-COSMOS, *XMM*-COSMOS and SXDS surveys were added to the 4 Ms CDFS sample described in § 2.2. Surveys of different depths and widths are essential to covering as homogeneously as possible the redshift-luminosity ($z - \mathrm{L_X}$) plane, since luminous QSOs are rare and a sizable number of them can be found just in wide and shallow surveys, while Seyfert-like objects require deep X-ray observations (and, hence, deep and pencil beam surveys) to be detected. Deep multi-wavelength coverage of the same fields is necessary to provide reliable (spectroscopic or photometric) redshift information for the largest possible number of X-ray sources, so that the final sample includes most of the high-redshift sources and as small as possible number of low-redshift interlopers. We will make use of a Maximum Likelihood procedure to fit a number of widely-adopted analytical models in literature to the HXLF in order to assess the evolution of the high-redshift AGN population. In § 3.1 we give a description of the X-ray surveys used in this chapter (besides the 4 Ms CDFS described in § 2.1), in § 3.2 we describe the sample of high-redshift AGN, in § 3.3 we derive the binned estimates of the HXLF, in § 3.4 we describe the evolutionary analytical models, in § 3.5 we report the fitting procedure and results and in § 3.5.1 we include a correction for redshift incompleteness. The number counts for the entire sample are presented in § 3.8. The findings are then discussed in § 3.9. All the results are included in Vito et al. (2014a).

3.1 X-ray surveys

3.1.1 *XMM*-COSMOS

The *XMM*-COSMOS survey (Hasinger et al. 2007) consists of 55 XMM pointings in the Cosmic evolution survey (COSMOS; Scoville et al. 2007) over an area of 2.13

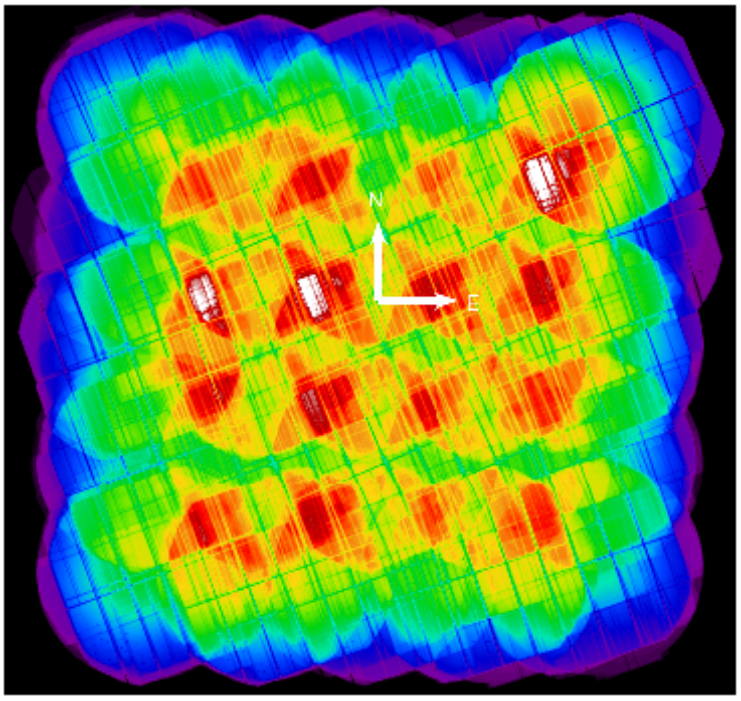

Figure 3.1: Soft band (0.5-2.0 keV) exposure map of *XMM*-COSMOS (from Cappelluti et al. 2009). The maximum exposure is ~ 84 ks (white), while the average exposure is ~ 68 ks (green).

deg^2, performed from December 2003 to June 2006. The European Photon Imaging Camera (EPIC) was used for a total integration time of ~ 1.55 Ms and an average exposure of ~ 70 ks. The X-ray source catalogue was presented by Cappelluti et al. (2009), while multi-wavelength identification was performed by Brusa et al. (2010).

Fig. 3.1 shows the soft-band $(0.5-2$ keV) exposure time of XMM-COSMOS. The sensitivity map is displayed in Fig. 3.2. The flux limit of the survey is $\sim 5\times10^{-16}/3\times10^{-15}/7\times10^{-15}$ in the $0.5-2/2-10/5-10$ keV band, respectively.

The *XMM*-COSMOS catalogue (Cappelluti et al. 2009) includes 1887 X-ray sources, 1621/1111/251 of which detected in the 0.5-2/2-10/5-10 keV band, respectively. A false colour X-ray image of the field is shown in Fig. 3.3. Multi-wavelength identification of 1797 X-ray point-sources (Brusa et al. 2010) took advantage of the Capak et al. (2007); Schinnerer et al. (2007); Bondi et al. (2008); Ilbert et al. (2009); Le Floc'h et al. (2009); McCracken et al. (2010) catalogues.

Ninety-nine per cent of the sources are associated to a multi-wavelength counterpart, while 17 sources remain with no optical identification. Spectroscopic redshifts were collected from Magellan/IMACS and MMT observation campaigns (Trump et al. 2007; 2009), VIMOS/zCOSMOS faint (Lilly et al. 2007) and bright (Lilly et al. 2009) projects, Sload Digital Sky Survey catalog (Adelman-McCarthy et al. 2006; Kauffmann et al. 2003), (Prescott et al. 2006) and Keck-II/Deimos observations (see Kartaltepe et al. 2010). Photometric redshifts were collected from Salvato et al. (2009). A Spectroscopic or photometric redshift is associated to almost 100% of the X-ray point-sources, with only 28 sources not having any redshift informa-

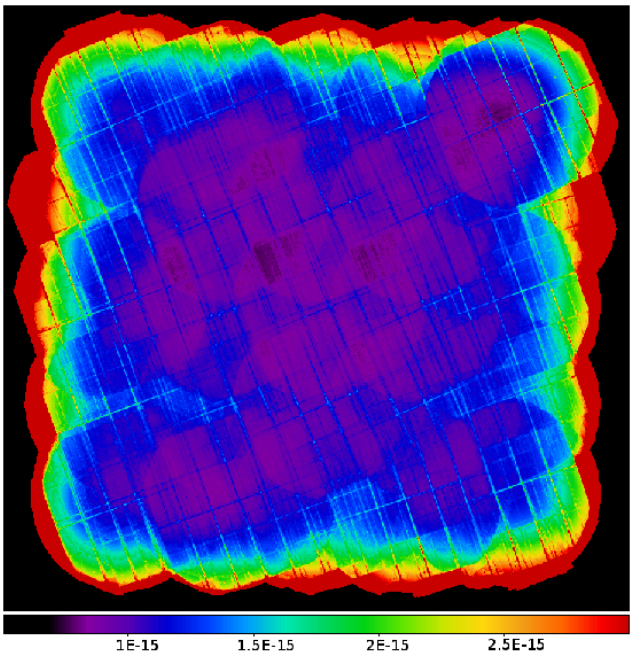

Figure 3.2: 0.5-2.0 keV sensitivity map of *XMM*-COSMOS (from Cappelluti et al. 2009), in units of erg $\mathrm{cm}^{-2}\mathrm{s}^{-1}$

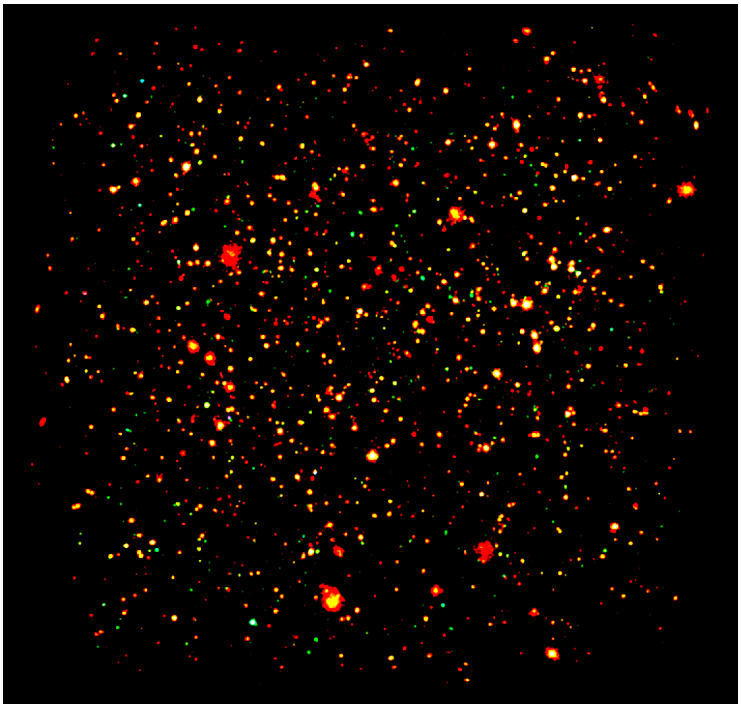

Figure 3.3: Composite X-ray image of the *XMM*-COSMOS field in the 0.5-2 (red), 2-4.5 (green) and 4.5-10 keV (blue) bands (from Cappelluti et al. 2009).

tion. Brusa et al. (2009) presented the optical and X-ray properties and the space density of a sample of 40 X-ray sources at $z > 3$.

Brusa et al. (2010) provide a classification of the X-ray sources with available optical spectra based on optical and X-ray information. In particular, a source is classified as Broad-Line AGN (BL AGN, 421 sources), if it shows at least one broad (FWHM$>$ 2000 km s^{-1}) emission line and Narrow-Line AGN (NL AGN, 370 sources) if it shows high-ionization emission lines or if the rest-frame, hard-band luminosity is larger than $2 \times 10^{42}\,\mathrm{erg\,s^{-1}}$. Other objects (53 sources) with available optical spectra are classified as normal galaxies, if emission-line ratios are consistent with star-forming galaxy emission or if only absorption lines are present in the spectra and the X-ray luminosity is $L_{2-10\ \mathrm{keV}} < 10^{42}\,\mathrm{erg\,s^{-1}}$ or they are undetected in the hard X-ray band. No classification is given for object lacking an optical spectrum.

3.1.2 *Chandra*-COSMOS

Chandra-COSMOS (Elvis et al. 2009) is a *Chandra* program consisting of 36 heavily-overlapping ACIS-I pointings of the central 0.9 deg^2 of the COSMOS field. Observations were carried on from November 2006 to June 2007. The total observing time is ~ 1.8 Ms, with an effective exposure of ~ 160 ks and ~ 80 ks in the inner 0.5 deg^2 and outer 0.4 deg^2, respectively.

The mosaic-like pattern of the survey is well visible in Fig. 3.4, which shows the exposure map in the soft-band of *Chandra*-COSMOS. The flux limit is $1.9/7.3/5.7 \times 10^{-16}\,\mathrm{erg\,cm^{-2}s^{-1}}$ in the $0.5-2/2-10/0.5-10$ keV band, respectively. The point-source catalogue (Elvis et al. 2009) includes 1761 X-ray sources, 1428/1182/1726 detected in the soft/hard/full band.

The multi-wavelength identification of the X-ray sources is presented by Civano et al. (2012), where *Chandra* sources were matched to the Capak et al. (2007); Sanders et al. (2007); McCracken et al. (2010) catalogues. Approximately 99% of the X-ray sources were matched to an optical/NIR counterpart, with an estimated spurious association $< 6\%$. Spectroscopic redshifts were collected from Magellan/IMACS and MMT observation campaigns (Trump et al. 2007; 2009), VIMOS/zCOSMOS faint (Lilly et al. 2007) and bright (Lilly et al. 2009) projects, SDSS DR7[1], (Prescott et al. 2006), Keck-II/Deimos observations (see Kartaltepe et al. 2010), Wright et al. (2010), van Dokkum et al. (2011) and Brammer et al. (2012). Photometric redshifts were collected from Salvato et al. (2011). Out of the 1761 X-ray detected sources, 1069 objects ($\sim 60\%$) have a spectroscopic redshift and 1693 objects have a photometric redshift ($\sim 96\%$). Civano et al. (2011) studied a sample of 81 $z > 3$ X-ray sources in *Chandra*-COSMOS and presented their number counts and space density.

Civano et al. (2012) classified the X-ray sources into spectroscopically BL AGN, non-BL objects (NL AGN or star-forming galaxies) and stars (see Tab. 3.1) on the basis of the presence or absence of broad emission-lines in the optical spectra. The classification from the SED fitting procedure (Salvato et al. 2011), based on the best-fitting templates, is also provided.

[1] http://classic.sdss.org/dr7/start/aboutdr7.html

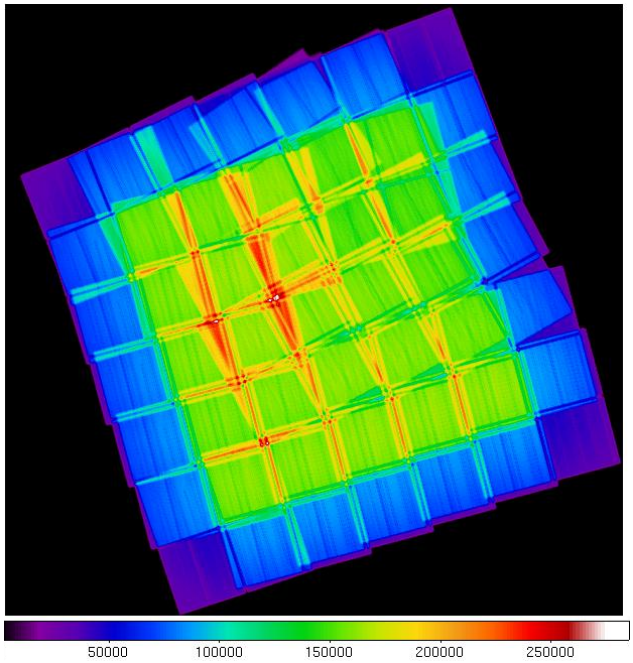

Figure 3.4: Soft band (0.5-2.0 keV) exposure map of *Chandra*-COSMOS (from Elvis et al. 2009). The coloured bar gives the exposure in units of seconds.

Table 3.1: Number of X-ray sources by spectral or photometric type (from Civano et al. 2012).

	Number of sources
Spectra	
Broad-Line AGN	316
Not Broad-Line AGN	534
Photo-z	
Unobscured AGN template	450
Obscured AGN template	104
Galaxy template	1101

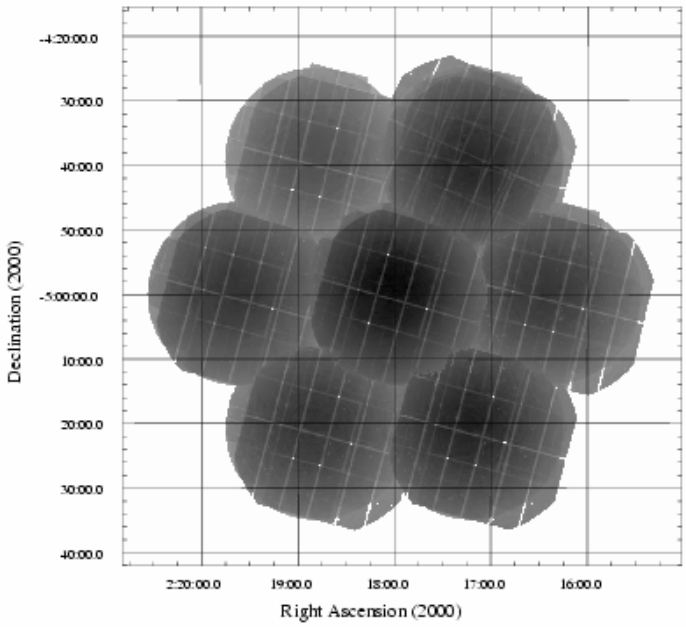

Figure 3.5: 0.5-4.5 keV sensitivity map of SXDS (from Ueda et al. 2008). Darker regions correspond to lower fluxes.

3.1.3 SXDS

The *SUBARU/XMM* Deep Survey (SXDS; Ueda et al. 2008) field was observed by *XMM* with seven pointings of the EPIC pn and MOS cameras between from July 2000 to January 2003. The X-ray survey was designed to cover almost entirely the optical field (Furusawa et al. 2008), for a total of ~ 1.14 deg^2.

Fig. 3.5 shows the $0.5-4.5$ keV sensitivity map. The flux limit of the SXDS is $6\times10^{-16}/8\times10^{-16}/3\times10^{-15}/5\times10^{-15}$ erg cm^{-2}s^{-1} in the $0.5-2/0.5-4.5/2-10/4.5-10$ keV band. The X-ray catalogue Ueda et al. (2008) contains 1245 sources, 866/1114/645/136 of which detected in the $0.5-2/0.5-4.5/2-10/4.5-10$ keV band. The X-ray sources are located in the field as shown by Fig. 3.6.

The optical identification of the X-ray sources, which took advantage of optical/NIR spectroscopic data taken by Subaru/Suprime-Cam and photometric data obtained by UKIRT/WFCAM and *Spitzer*/IRAC, at the time of this work is not yet published. Hiroi et al. (2012) reported that 781 out of the 866 soft-band detected sources fall into the Subaru/Suprime-Cam optical field and that 733 of them are AGN candidates. A spectroscopic redshift is available for 586 AGN candidates. Photometric redshifts were estimated by fitting up to 15 bands, from the far-UV to the mid-IR wavelengths, except for 6 objects detected in less than 4 optical bands. The redshift completeness is hence $\sim 99.2\%$. Hiroi et al. (2012) presented the space density and obscured fraction of a sample of 30 $z > 3$ AGN detected in the SXDS.

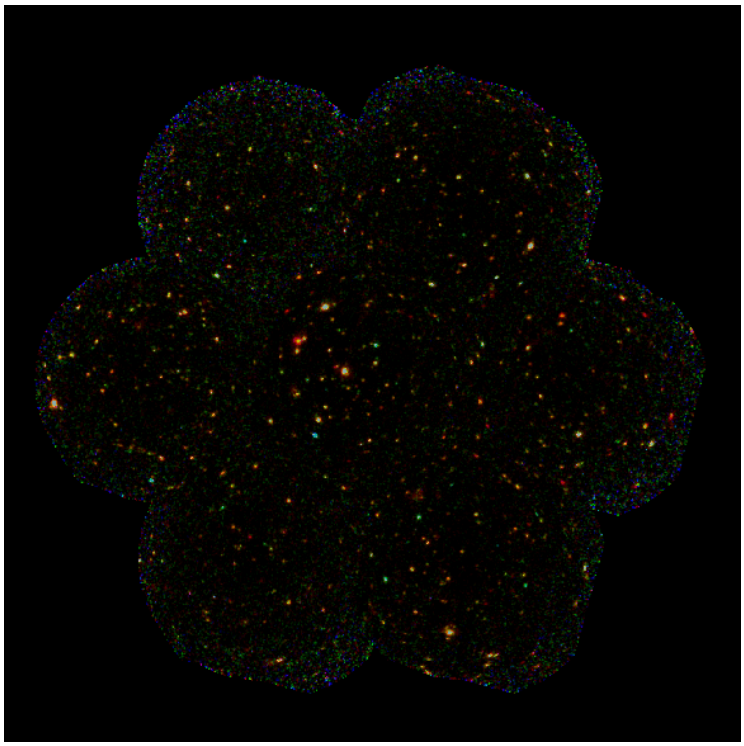

Figure 3.6: Composite X-ray image of the SXDS field in the 0.5-2 (red), 2-4.5 (green) and 4.5-10 keV (blue) bands(from Ueda et al. 2008).

3.2 The high-redshift sample

We collected a sample of 141 AGN detected in the soft X-ray band ($0.5-2$ keV) in the redshift range $3 < z < 5.1$. The $0.5-2$ keV band approximatively corresponds to the $2-10$ keV band at $z \approx 3$. The sample is assembled from four surveys of different sizes and depths. This is crucial to cover as homogeneously as possible the $L_X - z$ plane, since distant, Seyfert-like AGN can be detected only by the deepest pencil-beam surveys, while the rare, very luminous QSOs need wide (and hence shallower) surveys to be found. The choice of the fields was driven by the presence of massive multi-wavelength campaigns, which resulted in a very-high ($\gtrsim 95\%$) redshift completeness of the parent samples. We used both spectroscopic (78 objects, ~ 55 per cent of the entire sample) and photometric (63 objects, ~ 45 per cent) redshifts. In § 3.3 we will assign the sources to two different, absorption-based subsamples, where the threshold is set to $\log N_H = 23$. In the following sub-sections, we describe the samples of high-redshift AGN collected from each considered survey. We also quote the number of sources with no redshift information, which will be used in § 3.5.1 and the number of sources obscured by a column density $\log N_H > 23$. This information is summarised in Tab. 3.2.

3.2.1 The 4 Ms *Chandra* Deep Field South (CDF-S)

We used the sample of $z > 3$ X-ray selected AGN in the 4 Ms CDFS (see § 2.1) described by Vito et al. (2013) and in § 2.2, restricted to the redshift range $3 <$

$z < 5.1$ by discarding 3 objects which are associated to photometric redshifts $5.1 < z \lesssim 7.6$. Accounting only for the soft-band detected objects, the sample consists of 27 sources with intrinsic (i.e. rest-frame and corrected for absorption) luminosities $8 \times 10^{42} \lesssim \frac{L_{2-10\ \mathrm{keV}}}{\mathrm{erg\,s^{-1}}} \lesssim 10^{45}$. A spectroscopic redshift is available for 13 objects (~ 48 per cent), while a photometric redshift is associated to the remaining 14 sources. As for the redshift completeness, Vito et al. (2013) showed that 39 out of the 740 (~ 5.3 per cent) X-ray sources in the survey main catalogue (Xue et al. 2011) have neither a spectroscopic nor a photometric redshift. This number decreases to ~ 4.8 per cent (31 out of 650), if only the sources detected in the soft band are considered.

A spectral analysis was performed on this sample (see Vito et al. 2013 and § 2.3.2), assuming an absorbed power-law with photon index fixed to $\Gamma = 1.8$ and accounting for Galactic absorption ($N_H = 7 \times 10^{19}$ cm^{-2}). More complicated models could not be used because of the poor photon counting statistics characterising the sample. Sixteen out of 27 sources (~ 63 per cent) are highly obscured (best-fitting column density $N_H > 10^{23}$ cm^{-2}). The spectral parameters are reported in Tab. 3.3.

3.2.2 The Subaru/XMM-Newton Deep Survey (SXDS)

Hiroi et al. (2012) selected a sample of 30 AGN detected in the SXDS (see § 3.6) at $z > 3$ in the soft band, making use of the spectroscopic (20 objects) and photometric (10 objects) information available in that field.

In their analysis, the intrinsic spectrum is assumed to be a power-law with photon index fixed to $\Gamma = 1.9$ plus a contribution from a reflection component. We note that this shape mimics a spectrum with an effective photon index of $\Gamma \approx 1.8$. The sample lies in the redshift range $3 < z \leq 4.5$ and is characterised by intrinsic luminosities in the range $10^{44} \lesssim \frac{L_{2-10\ \mathrm{keV}}}{\mathrm{erg\,s^{-1}}} \lesssim 10^{45}$. For each source, they derived an estimate of the column density by mean of the hardness ratio (HR):

$$\mathrm{HR} = \frac{\mathrm{H} - \mathrm{S}}{\mathrm{H} + \mathrm{S}} \tag{3.1}$$

where S and H are the count rates in the $0.3 - 1.0$ and $1.0 - 4.5$ keV band, respectively. One source is obscured by a column density $N_H > 10^{23}$ cm^{-2}. We used the Hiroi et al. (2012) sample and spectral parameters (shown in Tab. 3.3). Six X-ray sources in this survey have no redshift information.

3.2.3 The *Chandra* Cosmological Evolution Survey (C-COSMOS)

Using the new redshifts available in this field (Salvato et al. 2011; Civano et al. 2012, Lilly et al. in prep), we refined the high-redshift sample in *Chandra*-COSMOS (see § 3.1.2) presented by Civano et al. (2011) and found 62 soft-band detected objects at $3 < z < 5.1$. Accounting only for the soft-band selected objects, the Civano et al. (2011) sample consisted in 65 $z > 3$ objects. The difference is the result of the exclusion of 7 sources (ID 293, 666, 871, 1112, 1276, 1304, and 11633) of the Civano et al. (2011) sample with a new, $z < 3$ redshift in Civano et al. (2012), and the inclusion of 6 objects (ID 529, 688, 1311, 1392, 2059 and 3449), not present in Civano et al. (2011), but identified as high-redshift AGN in Civano et al. (2012). *Chandra* ID 125 and 2550 have a redshift $z > 5.1$ and therefore will not be considered in this work.

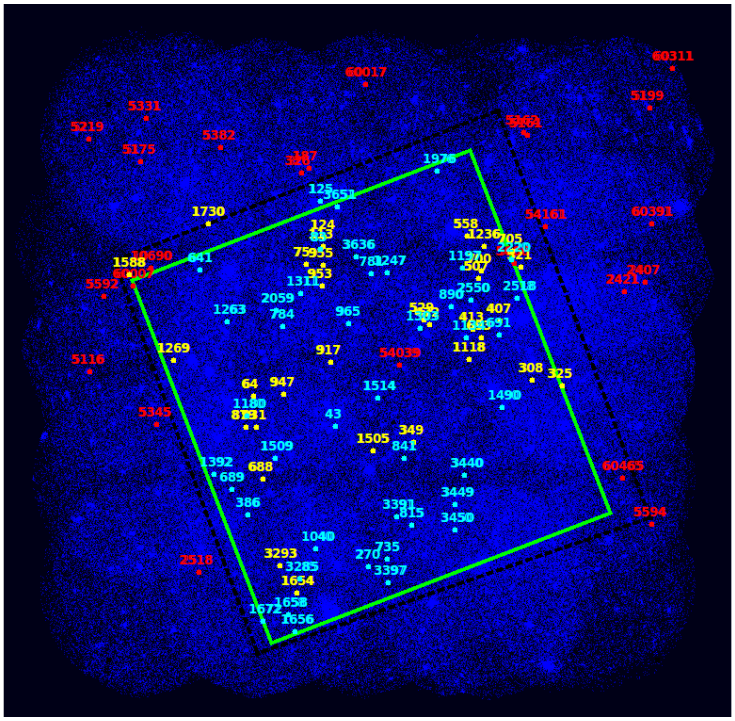

Figure 3.7: Soft-band (0.5 − 2 keV) *XMM*-COSMOS image. Black dashed rectangle represents the original *Chandra*-COSMOS field, while the green rectangle is the reduced *Chandra* field used in this work (see § 3.2.3). Cyan, red and yellow points are the high-redshift sources detected only by *Chandra*, only by *XMM*-Newton and by both satellites, respectively. The IDs are the same than in Tab. 3.3.

Almost half of our sample (30 out of 62 AGN) was also detected by the XMM-Newton survey in the COSMOS field, which encompasses the *Chandra* observations. Since in the outskirts of the *Chandra* field (black rectangle in Fig. 3.2.3) the sensitivity drops below the sensitivity of XMM-COSMOS, there are a number of sources detected by XMM-Newton whose flux is below the *Chandra* limit at that point. In order 1) to count each source only once and 2) not to lose the AGN detected only in XMM-COSMOS, we did not consider all the C-COSMOS field, but only the central region. The proper way to define it would be the region where *Chandra* is more sensitive than XMM-Newton. However, for simplicity during the evaluation of the sky-coverage (see § 3.3.1), we defined the new C-COSMOS area by looking at the overlapping XMM and *Chandra* images and cutting off an external frame of the *Chandra* field as narrow as possible, with an area of ~ 0.20 deg^2, to fulfil the two requirements mentioned above.

All the sources detected either only by *Chandra* (cyan points in Fig. 3.2.3) or by both *Chandra* and XMM-Newton (yellow points) in the resulting 0.72 deg^2 region (green rectangle) will then be counted in the C-COSMOS sample. Three objects (*Chandra* ID 521, 1730 and 325) detected by both *Chandra* and XMM-Newton, fall outside the newly defined boundaries of C-COSMOS; therefore, they are counted in the XMM-COSMOS sample (XMM ID 5120, 5259 and 5347, respectively).

Hence, the final high-redshift C-COSMOS sample consists of 59 AGN in the

redshift range $3 < z \leq 5.1$. A spectroscopic redshift is associated to 30 of them. We performed a spectral analysis assuming the same spectral model described in § 2.3.2 (i.e. an absorbed power-law with $\Gamma = 1.8$ and Galactic column density of 2.5×10^{20} cm^{-2}; Kalberla et al. 2005) and fitted the spectra of the 59 sources using the Cash statistics (Cash 1979) to estimate the best-fitting parameters (Tab. 3.3). We derived intrinsic luminosities of $10^{43} \lesssim \frac{L_{2-10\ \mathrm{keV}}}{\mathrm{erg\,s^{-1}}} \lesssim 2 \times 10^{45}$ and a fraction of ~ 32 per cent of very obscured sources ($N_H > 10^{23}$ cm^{-2}). The poor spectral quality of *Chandra* ID 3449 prevented us from investigating the effect of absorption, hence a simple power-law was assumed as spectral model in order to derive the intrinsic hard band luminosity. This source will be conservatively considered unobscured in the following sections.

3.2.4 The XMM-Newton Cosmological Evolution Survey (XMM-COSMOS)

As for the *XMM*-COSMOS survey (see § 3.1.1), we considered as parent sample the 1797 sources reported in the Brusa et al. (2010) multi-wavelength catalogue. Redshift were collected from Brusa et al. (2010), Salvato et al. (2011) and Civano et al. (2012) and 53 soft-band detected objects resulted to be at $z > 3$.

Brusa et al. (2009) presented results on a sample of high-redshift AGN in XMM-COSMOS. In particular, they selected 39 soft-band detected AGN $z > 3$ with $F_{0.5-2\ \mathrm{keV}} > 10^{-15}$ erg cm^{-2}s^{-1}. Because of the new redshift information we collected, these two samples have 34 sources in common.

Once the central reduced C-COSMOS field is cut out (see Fig. 3.2.3), the remaining XMM-COSMOS area is ~ 1.45 deg^2 and the number of soft-detected object that are placed in that frame are 930, with a redshift completeness of 99.4 per cent (i.e. 6 objects have no redshift information). Twenty-five objects are high-redshift sources (19 of which are included in the Brusa et al. (2009) sample); 15 of them have a spectroscopic redshift. All but one of the remaining 28 high-redshift AGN, which fall on the reduced C-COSMOS area, are detected by Chandra and are counted in the C-COSMOS sample (see § 3.2.3)[2].

Mainieri et al. (2007; 2011) performed a detailed spectral analysis on a subsample of X-ray sources in XMM-COSMOS, including 16 objects in our high-redshift sample. They assumed an absorbed power-law with photon index free to vary. Only one of them resulted to be heavily obscured (XID 2518, $\log \mathrm{N_H} = 23.07^{+0.35}_{-0.37}$), with a very steep photon index ($\Gamma > 2$). Since the photon counting statistics was too poor to perform a spectral analysis on all these objects and in order to be consistent with the spectral assumptions used for the other surveys and with the procedure adopted in § 3.3, we conservatively consider all the sources in our $z > 3$ sample as unabsorbed AGN. This is also justified since, at similar fluxes, the SXDS sample (see § 3.2.2) included only unabsorbed object, with just one exception. We derived the rest-frame hard-band luminosities from the soft-band flux reported in Cappelluti

[2]XID 54039 is a high-redshift X-ray source which falls on the central, reduced Chandra field, but is detected only by XMM. This is probably due to statistical fluctuations affecting the flux of the source or the nearby background, or even to intrinsic variability of the source. Since in the reduced C-COSMOS region we considered only sources detected by *Chandra*, we excluded it from the high-redshift sample.

Table 3.2: Main properties of the individual surveys

FIELD (1)	A [deg^2] (2)	% $_{\rm compl}$ (3)	N (N$_{\rm abs}$) (4)	% $_{\rm zspec}$ (5)
CDF-S	0.129	95	27 (16)	48
C-COSMOS$_{red}$	0.717	98	59 (19)	51
XMM-COSMOS$_{red}$	1.453	99	25 (0)	60
SXDS	1.006	99	30 (1)	67
TOTAL	3.305	98	141 (36)	55

(1) Field; (2) nominal area; (3) redshift completeness of the soft-band detected parent sample; (4) number of soft-band detected AGN at $3 < z < 5.1$, the number of sources included in the A subsample ($\log N_{\rm H} > 23$; see § 3.3.1) is between brackets; (5) fraction of high-redshift AGN with a spectroscopic redshift. These numbers refer to the reduced field for *Chandra* and XMM-COSMOS and to the overlapping region between the X-ray and optical surveys for the SXDS field (see § 3.2 and § 3.6).

et al. (2009), assuming a simple power-law with $\Gamma = 1.8$ (Tab. 3.3).

3.2.5 General properties of the sample

Fig. 3.8 shows the redshift distribution of the sample. The column density is plotted against redshift (left panel) and luminosity (right panel) for each source of the sample in Fig. 3.9. We defined the column density to be constrained if its lower limit at the 90 per cent confidence level (c.l.) is larger than zero, otherwise we plot its 90 per cent c.l. upper limit as a downward pointing arrow. Since in Hiroi et al. (2012) errors are quoted at the 1σ c.l., for sources detected by the SXDS survey we applied the same definition, but changed the confidence limit to the 90 per cent c.l. Thirty-six out of the 141 objects are obscured by a column density $N_H > 10^{23}$ cm^{-2}. Most of them are detected by the CDFS and C-COSMOS survey, as expected, since they are the two deepest surveys among those considered in this work.

Fig. 3.10 shows the luminosity plotted against the redshift. The importance of collecting samples from different (in terms of deepness and area) surveys to sample different region in the $L_X - z - N_H$ space can be inferred from these figures.

3.3 The binned HXLF

The differential luminosity function of any population of extra-galactic objects can be defined as the number N of objects per unit comoving volume V per unit logarithmic luminosity $\log L_{\rm X}$

$$\phi = \frac{{\rm d}\Phi}{{\rm dlog}L_{\rm X}}(z,\ \log L_{\rm X}) = \frac{{\rm d}^2 N}{{\rm d}V\ {\rm dlog}L_{\rm X}} \tag{3.2}$$

A binned representation of the HXLF can be derived using the $1/V_{max}$ method (Schmidt 1968; Avni & Bahcall 1980), but it suffers from systematic bias for objects close to the flux limit of the surveys (Page & Carrera 2000). Since this is the case

Table 3.3: Main information and spectral parameters of high-redshift sample.

ID	RA	DEC	zadopt	ztype	zref	N_H	$F_{0.5-2\ \mathrm{keV}}$	$L_{2-10\ \mathrm{keV}}$
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
				C-COSMOS				
43	150.18087	2.075997	3.01	1	1	< 3	1.41	1.49
64	150.36472	2.143831	3.328	1	1	< 7	3.12	4.80
75	150.24770	2.442225	3.029	1	1	22^{+9}_{-7}	2.57	8.35
83	150.21418	2.475118	3.075	2	1	17^{+9}_{-7}	2.00	5.52
113	150.20884	2.482010	3.333	1	1	< 4	3.04	4.03
124	150.20536	2.502848	3.072	2	1	13^{+15}_{-10}	1.71	4.14
270	150.10734	1.759256	4.16	1	1	< 72	0.70	1.84
308	149.73615	2.179954	4.255	1	1	< 30	1.24	4.05
349	150.00438	2.038978	3.515	1	1	< 5	1.62	2.43
386	150.37885	1.876099	3.33	2	1	21^{+33}_{-19}	0.63	2.19
407	149.80849	2.313858	3.471	2	1	< 4	2.67	3.88
413	149.86968	2.294064	3.345	1	1	21^{+10}_{-8}	2.71	9.05
472	149.96920	2.304833	3.155	1	1	17^{+16}_{-10}	1.24	3.57
507	149.85845	2.409299	4.108	2	1	15^{+8}_{-7}	4.89	18.50
529	149.98156	2.315056	3.017	1	2	17^{+9}_{-7}	1.93	5.13
558	149.88247	2.505174	3.1	2	1	4^{+4}_{-3}	4.34	6.90
688	150.34505	1.958014	3.065	2	1	< 23	0.84	1.66
689	150.41521	1.934286	3.681	2	1	< 16	0.86	1.43
691	149.81217	2.282920	3.297	2	1	< 6	0.90	1.16
693	149.85148	2.276539	3.371	1	1	< 7	1.35	1.85
700	149.85151	2.426858	3.35	2	1	46^{+33}_{-23}	5.43	3.52
781	150.10093	2.419495	4.66	1	1	< 61	0.74	4.14
784	150.30071	2.300688	3.498	1	1	< 76	0.33	1.21
815	150.00937	1.852672	4.032	2	1	< 21	0.71	1.45
879	150.38347	2.074682	3.859	1	2	76^{+132}_{-70}	0.16	1.96
890	149.91958	2.345473	3.021	1	1	< 14	1.06	1.19
917	150.19256	2.219909	3.09	1	1	< 20	0.87	1.76
931	150.35964	2.073574	4.917	1	1	< 43	0.62	2.39
947	150.29719	2.148829	3.328	1	1	16^{+29}_{-12}	0.56	1.65
953	150.21070	2.391473	3.095	1	1	< 22	0.90	1.37
955	150.20899	2.438581	3.715	1	1	< 6	1.05	1.78
965	150.15215	2.307818	3.175	1	1	104^{+50}_{-42}	0.26	5.39
1040	150.22593	1.799779	3.264	2	1	52^{+37}_{-27}	0.49	3.63
1118	149.87920	2.225811	3.65	1	1	< 13	1.72	3.68
1197	149.89429	2.433144	3.382	1	1	< 12	0.27	0.37
1236	149.84572	2.481628	3.375	2	1	< 28	0.74	1.20
1263	150.42519	2.312089	3.092	2	1	59^{+70}_{-40}	0.24	2.03
1269	150.54647	2.224128	3.506	2	1	12^{+12}_{-9}	2.25	6.07
1303	149.99044	2.297347	3.026	1	1	< 24	0.82	1.43
1311	150.25977	2.376108	3.717	1	1	< 35	0.67	1.37
1392	150.45489	1.967361	3.485	1	1	< 47	0.72	1.73
1490	149.80460	2.118866	3.791	2	1	< 52	0.34	1.02
1505	150.09688	2.021498	3.546	1	1	< 20	0.69	1.05
1509	150.31788	2.004926	3.428	2	1	< 40	0.58	1.21
1514	150.08617	2.138865	5.045	2	1	94^{+73}_{-55}	0.30	5.14
1654	150.26739	1.700929	3.412	2	1	< 11	0.96	1.25
1656	150.27158	1.613616	3.466	2	1	32^{+35}_{-28}	1.26	6.10

(1) source identification number as in Elvis et al. (2009),Xue et al. (2011), Ueda et al. (2008) and Cappelluti et al. (2009) for the CDFS, C-COSMOS, SXDS and XMM-COSMOS sample, respectively; (2) right ascension and (3) declination (J2000) of the X-ray source; (4) adopted redshift; (5) 1: spectroscopic redshift, 2: photometric redshift; (6) redshift reference; 1: Civano et al. (2012), 2: zCOSMOS, 3: Vito et al. (2013), 4: Hiroi et al. (2012), 5: Brusa et al. (2010), 6: Salvato et al. (2011); (7) best-fitting column density in units of $[10^{22}\ \mathrm{cm}^{-2}]$, errors are at the 90 per cent confidence limit; (8) soft-band flux in units of $[10^{-15}\ \mathrm{erg\,cm^{-2}s^{-1}}]$; (9) intrinsic (i.e. absorption-corrected) rest-frame $2-10$ keV luminosity, in units of $10^{44}\ \mathrm{erg\,s^{-1}}$.

Table 3.3: Continued.

ID	RA	DEC	zadopt	ztype	zref	N_H	$F_{0.5-2\ \mathrm{keV}}$	$L_{2-10\ \mathrm{keV}}$
				C-COSMOS				
1658	150.28776	1.650687	3.871	2	1	< 10	1.64	3.05
1672	150.34429	1.635908	3.805	2	1	< 67	1.19	4.87
2059	150.31583	2.336873	4.216	2	1	127^{+165}_{-99}	0.11	2.96
2220	149.78373	2.452045	5.07	1	1	< 90	0.80	4.68
2518	149.77105	2.365836	3.447	2	1	< 118	0.24	1.45
3293	150.30590	1.761653	3.265	2	1	< 46	0.38	0.96
3391	150.04273	1.872178	3.371	1	1	< 12	0.81	2.12
3397	150.06217	1.722708	3.033	2	1	< 30	0.96	2.02
3440	149.88939	1.966029	3.053	2	1	< 28	0.60	0.65
3449	149.91073	1.899629	3.063	2	1	-1	0.11	0.12
3636	150.13372	2.457497	3.189	1	1	25^{+39}_{-22}	0.50	1.83
3651	150.17640	2.569708	3.144	2	1	< 35	0.50	0.58
				CDFS				
27	52.96054	-27.87706	4.385	2	3	68^{+41}_{-29}	0.25	2.80
100	53.01658	-27.74489	3.877	2	3	38^{+27}_{-24}	0.15	0.89
107	53.01975	-27.66267	3.808	2	3	< 71	0.48	1.52
132	53.03071	-27.82836	3.528	2	3	13^{+16}_{-11}	0.11	0.31
170	53.04746	-27.87047	3.999	2	3	35^{+4}_{-4}	0.16	9.43
235	53.07029	-27.84564	3.712	1	3	< 36	0.04	0.08
262	53.07854	-27.85992	3.66	1	3	85^{+22}_{-20}	0.12	1.63
283	53.08467	-27.70811	3.204	2	3	55^{+39}_{-22}	0.10	0.80
285	53.08558	-27.85822	4.253	2	3	< 6	0.06	0.12
331	53.10271	-27.86061	3.78	2	3	< 23	0.05	0.08
371	53.11158	-27.76789	3.101	2	3	35^{+12}_{-10}	0.12	0.57
386	53.11796	-27.73439	3.256	1	3	< 7	0.09	0.11
412	53.12442	-27.85169	3.7	1	3	82^{+12}_{-11}	0.21	2.85
458	53.13854	-27.82128	3.474	1	3	93^{+74}_{-70}	0.02	0.26
521	53.15850	-27.73372	3.417	2	3	< 23	0.09	0.21
528	53.16158	-27.85606	3.951	1	3	74^{+24}_{-24}	0.10	1.16
546	53.16533	-27.81419	3.064	1	3	52^{+4}_{-4}	0.67	4.73
556	53.17012	-27.92975	3.528	2	3	97^{+9}_{-10}	0.76	11.40
563	53.17442	-27.86742	3.61	1	3	6^{+2}_{-2}	2.07	4.45
573	53.17850	-27.78411	3.193	1	3	3^{+2}_{-2}	0.81	1.19
588	53.18467	-27.88103	3.471	1	3	< 3	0.65	0.84
642	53.20821	-27.74994	3.769	2	3	< 13	0.09	0.15
645	53.20933	-27.88119	3.47	1	3	15^{+2}_{-2}	3.16	9.24
651	53.21529	-27.87033	4.658	2	3	151^{+39}_{-35}	0.13	4.64
674	53.24004	-27.76361	3.082	1	3	< 7	0.70	0.86
700	53.26250	-27.86308	4.253	2	3	18^{+16}_{-14}	0.74	3.11
717	53.28000	-27.79892	4.635	1	3	87^{+65}_{-51}	0.12	1.85
				SXDS				
16	33.93335	-4.92384	3.512	1	4	< 7	3.16	3.57
99	34.08598	-5.28820	3.19	1	4	< 1	4.68	4.19
154	34.14188	-4.90640	3.6	2	4	4^{+10}_{-3}	1.58	1.89
177	34.16076	-5.17883	3.182	1	4	< 4	3.35	2.98
179	34.16469	-4.72107	3.426	2	4	< 1	1.23	1.31
284	34.24511	-4.81274	3.046	2	4	< 3	2.44	1.96
287	34.24634	-4.83036	4.09	2	4	< 4	5.12	8.30
335	34.27461	-5.22714	3.222	1	4	8^{+13}_{-7}	2.60	2.39

Table 3.3: Continued.

ID	RA	DEC	zadopt	ztype	zref	N_H	$F_{0.5-2\ \mathrm{keV}}$	$L_{2-10\ \mathrm{keV}}$
				SXDS				
342	34.28106	-4.56844	3.048	2	4	25^{+120}_{-19}	6.57	5.28
385	34.32280	-5.08699	3.39	2	4	< 8	2.66	2.76
422	34.34430	-5.39438	3.422	1	4	3^{+2}_{-2}	6.02	6.38
449	34.35595	-4.98441	3.328	1	4	< 4	2.58	2.56
459	34.36505	-5.28871	3.969	1	4	< 1	1.60	2.42
489	34.38058	-5.09214	3.334	2	4	7^{+5}_{-4}	4.65	4.63
508	34.39336	-5.08771	3.975	1	4	< 7	2.27	3.45
520	34.39922	-4.70872	3.292	1	4	< 15	1.57	1.52
564	34.43107	-4.54404	3.204	1	4	< 6	3.01	2.72
650	34.49477	-5.13762	3.032	1	4	< 1	2.18	1.73
700	34.52910	-5.06942	3.128	1	4	5^{+2}_{-2}	7.22	6.18
742	34.56485	-5.40081	3.114	1	4	< 2	6.84	5.78
788	34.59990	-5.10029	4.096	2	4	< 6	2.07	3.37
809	34.61807	-5.26411	3.857	1	4	< 6	1.96	2.76
824	34.63105	-4.73291	3.699	1	4	< 18	2.95	3.77
835	34.64068	-5.28748	3.553	1	4	< 3	1.34	1.55
888	34.68521	-4.80691	4.55	1	4	< 17	3.72	7.76
904	34.69847	-5.38866	3.02	1	4	< 1	21.60	1.69
926	34.72080	-5.01810	3.264	1	4	10^{+11}_{-6}	1.09	10.40
930	34.72563	-5.52342	3.49	2	4	< 4	4.67	5.19
1032	34.80911	-5.17238	3.584	2	4	< 6	4.19	4.96
1238	35.09164	-5.07500	4.174	1	4	< 2	2.60	4.42
				XMM-COSMOS				
187	150.240869	2.658730	3.356	1	5		3.80	4.85
326	150.256847	2.646315	3.003	1	5		1.20	1.19
2421	149.528973	2.380177	3.097	1	5		5.20	5.53
2518	150.489739	1.746145	3.176	1	5		1.30	1.46
5116	150.735835	2.199860	3.5	1	1		4.10	5.75
5120	149.761660	2.434940	3.647	2	6		2.50	3.84
5161	149.748141	2.732395	3.169	2	5		1.60	1.79
5162	149.755002	2.738672	3.524	1	5		1.90	2.70
5175	150.620110	2.671575	3.143	1	5		6.60	7.26
5199	149.471980	2.793812	3.626	1	5		2.70	4.10
5219	150.736810	2.722455	3.302	1	5		2.20	2.70
5259	150.466760	2.531320	4.45	2	5		1.10	2.63
5331	150.608403	2.769818	3.038	1	5		4.60	4.68
5345	150.584647	2.081176	3.296	2	6		1.90	2.33
5347	149.669389	2.167760	3.089	1	1		2.00	2.11
5382	150.440006	2.703517	3.465	1	5		1.10	1.51
5592	150.703941	2.369606	3.749	1	5		2.20	3.60
5594	149.467412	1.855175	4.161	1	5		1.30	2.68
5606	149.776790	2.444055	4.166	1	5		1.30	2.69
10690	150.596966	2.432707	3.1	2	5		1.30	1.39
54161	149.707096	2.525978	3.003	2	6		0.80	0.79
60017	150.112295	2.845933	3.062	2	6		1.30	1.35
60311	149.419259	2.883111	3.329	2	6		3.00	3.76
60391	149.467610	2.531345	3.518	2	6		1.00	1.42
60465	149.532828	1.958666	3.146	2	6		1.00	1.10

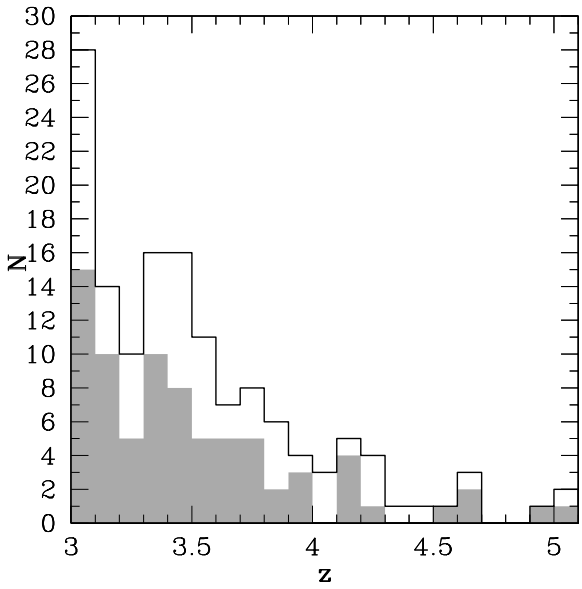

Figure 3.8: Redshift distribution of the 141 sources in the redshift range $3 < z < 5.1$ (black histogram) and the fraction of sources with spectroscopic redshift (grey shaded histogram)

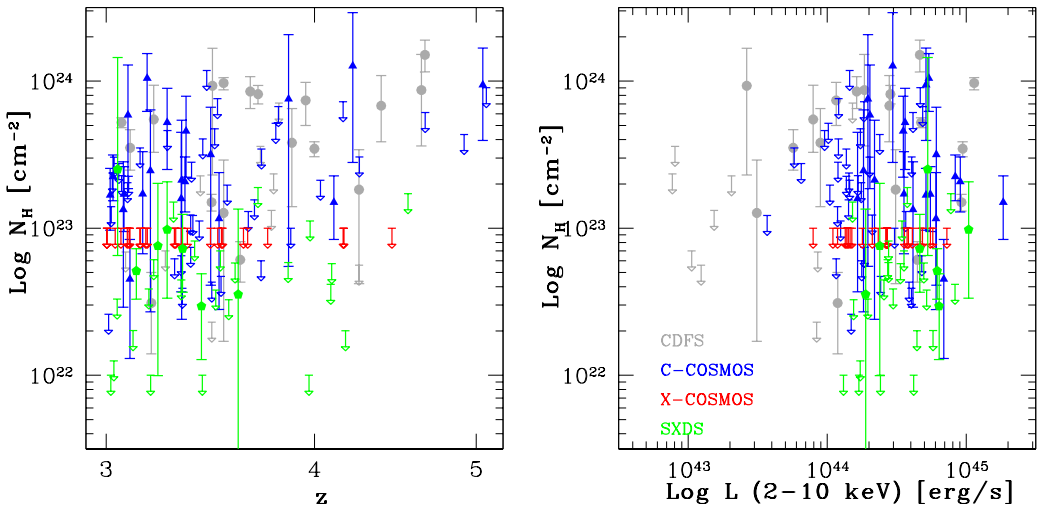

Figure 3.9: Column density as a function of redshift (left panel) and rest-frame, intrinsic hard-band luminosity (right panel) for the 141 sources at $3 < z < 5.1$. Grey, blue, red and green symbols represent sources from the CDFS, C-COSMOS, XMM-COSMOS and SXDS sample, respectively. Errors are plotted at the 90 per cent confidence level (c.l.). XMM-COSMOS sources are assumed to be unobscured and are plotted as upper limits corresponding to $N_H < 10^{23}$ cm^{-2}.

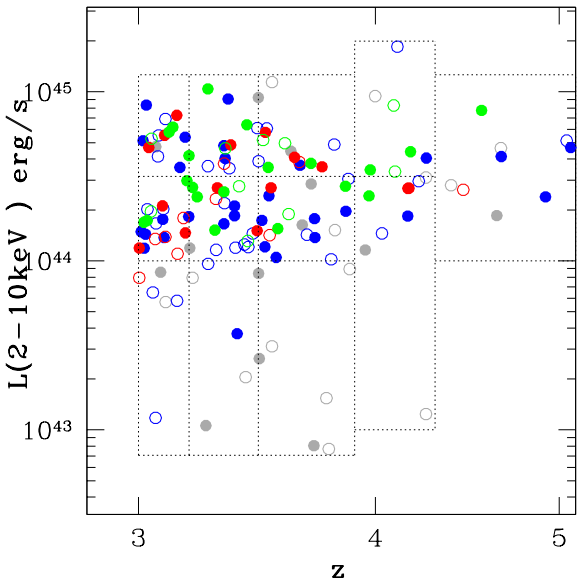

Figure 3.10: Rest-frame, intrinsic hard-band luminosity as a function of redshift. The colour code is the same as in Fig. 3.9. Filled and empty circles refer to sources with spectroscopic and photometric redshift, respectively. The dashed rectangles represent the luminosity-redshift bins considered in § 3.3.1.

of most sources included in our high-redshift sample, we preferred to use the Page & Carrera (2000) method, which was developed to fix this problem. Therefore, our estimation of the binned HXLF in a given redshift (Δz) and luminosity ($\Delta \mathrm{logL_X}$) bin has the following form:

$$\phi = \frac{N}{\iint \Omega \Theta \frac{\mathrm{d}V}{\mathrm{d}z} \; \mathrm{d}z \; \mathrm{dlogL_X}} \tag{3.3}$$

The Page & Carrera (2000) method assumes that the luminosity function variation inside a bin $\Delta \mathrm{logL_X} - \Delta z$ is small enough that can be neglected. This assumption is not necessarily true in our case, since we will make use of relatively large bins of redshift and luminosity. However, this would not affect the fit parameters which will be derived in § 3.5, since the fit will be performed on unbinned data. Moreover, this method provides a formally correct derivation of the errors on the binned HXLF, while the $1/V_{max}$ method does not.

3.3.1 Sky coverage and absorption

The $0.5-2$ keV sky coverage of the CDFS survey was taken from Xue et al. (2011). As for the SXDS survey, we used the sky coverage of the overlapping region between the X-ray and optical surveys (Ueda Y., private communication), used also by Hiroi et al. (2012) and derived from the curve presented in Ueda et al. (2008). The C-COSMOS sky coverage was published by Elvis et al. (2009). Since we cut out an external frame of this field (see § 3.2.3), a region of low-medium exposure, we imposed the maximum area of the survey to be 0.72 deg^2, the area of the reduced

C-COSMOS field, without any change in the normalization and shape of the sky coverage at low fluxes. As for the reduced XMM-COSMOS field, we re-computed the sky coverage as in Cappelluti et al. (2009), excluding the central 0.72 deg^2 region (see § 3.2.4).

Two main issues arise when considering the sky coverage curves: 1) the sky coverage for each survey was computed assuming a different spectral shape (Xue et al. 2011; Elvis et al. 2009; Cappelluti et al. 2009; Ueda et al. 2008); 2) sources with a given hard-band, intrinsic luminosity, have different soft-band fluxes and spectral shape, because of the effect of absorption. To address these issues, we converted the sky coverage as a function of flux of the i-th survey $\Omega_i(F)$ into a sky coverage as a function of count rate $\Omega_i(CR)$, according to the specific spectral shape assumed for each reference paper. Therefore, the products are no more dependent on a specific spectral shape.

Then, the full sample was splitted between two subsamples: sources with a column density (column 8 of Tab. 3.3) $\log N_H > 23$ were included in the subsample A (absorbed; 36 objects), the others were counted in the subsample U (unabsorbed; 105 objects). We used this threshold because it is approximately the N_H detection limit at $z \gtrsim 3$: lower values of column density do not affect significantly the observed soft-band spectrum of a X-ray source, since the photoelectric cut-off shifts to lower energies (e.g. Vito et al. 2013) and, hence, cannot be constrained.

The $\Omega_i(CR)$ are re-converted into sky coverage as a function of flux assuming as spectral shape a simple power-law with spectral index $\Gamma = 1.8$ and no absorption for the subsample U ($\Omega_i{}^{\rm U}$) and accounting for an absorption due to a column density $\log N_H = 23.5$ at $z = 3.5$ (approximatively the median values of N_H and redshift for obscured sources, see Tab. 3.3) for the subsample A ($\Omega_i{}^{\rm A}$). The use of the median N_H and z is justified by the need of a single spectral shape during the computation of the binned HXLF and by the large uncertainties on the column density of each individual source. All luminosities for the CDFS, C-COSMOS and XMM-COSMOS samples were computed with the above photon index. The luminosity for the SXDS sample were obtained by Hiroi et al. (2012) assuming a power-law with $\Gamma = 1.9$ plus a reflection component. However, this difference in the spectral shape is completely negligible for the final results.

In this procedure, we used typical (i.e. on axis) response matrices of the CDFS, C-COSMOS and XMM-COSMOS surveys to compute the energy-to-count-rate conversion factors (ECF) for the relative samples. The adopted matrices do not affect significantly the final results (using different responses resulted in a difference of < 5 per cent), since the conversion is made twice in the reverse order (from flux to count rate and then from count rate to flux) with the same matrix. As for the SXDS survey, we used the ECF provided by Ueda et al. (2008) for different values of effective photon index. We could use the ECF corresponding to $\Gamma = 1.8$ for the U subsample. For the A subsample an effective photon index $\Gamma \sim -1$ in the soft band was derived from an intrinsic $\Gamma = 1.8$ power-law spectrum absorbed by a column density $\log N_H = 23.5$ at $z = 3.5$. Since Ueda et al. (2008) did not provide the ECF corresponding to $\Gamma = -1$, we linearly interpolated the values they reported and derived the ECF.

We computed separately the HXLF for the subsamples U ($\phi^{\rm U}$) and A ($\phi^{\rm A}$). The sky coverage $\Omega^{\rm U} = \sum_i \Omega_i{}^{\rm U}$ (Fig. 3.11, upper panel) is used to compute $\phi^{\rm U}$

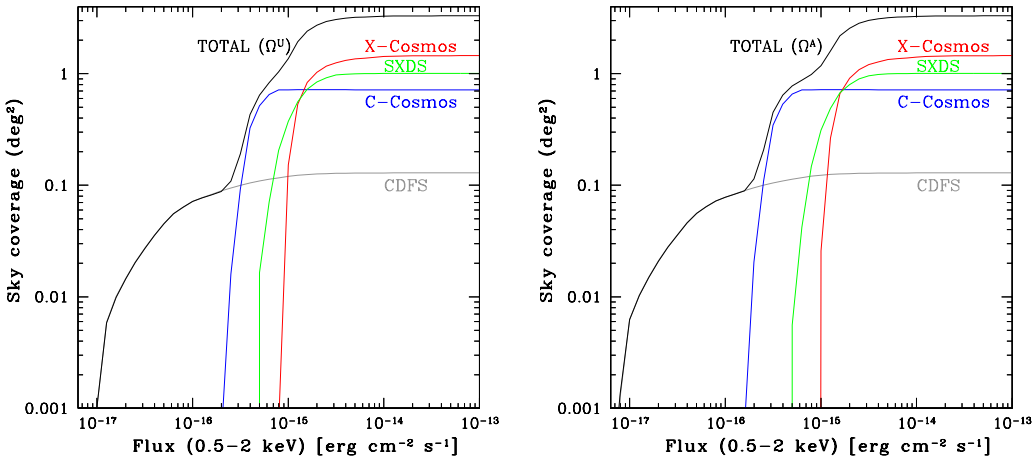

Figure 3.11: Sky coverage used for the computation of the HXLF for the unabsorbed (*left panel*) and absorbed (*right panel*) source sample. The total sky coverage (black line) is computed as the sum of the individual surveys sky coverage, as described in § 3.3.1.

and, similarly, $\Omega^{\rm A} = \sum_i \Omega_i{}^{\rm A}$ (Fig. 3.11, lower panel) to compute $\phi^{\rm A}$. Using this formalism, we adopted the coherent addition of samples in the sense given by Avni & Bahcall (1980, i.e. the proper addition of samples from different flux-limited surveys, assuming that each object could in principle be detected in all of them) and already used in a similar way by e.g. Yencho et al. (2009). Though the curves in Fig. 3.11 appear very similar, we note that close to the flux limit of a survey, the coverages for the U and A subsamples can be significantly different (e.g., a factor of ~ 4 for C-COSMOS at $F_{0.5-2\ \rm keV} = 3 \times 10^{-16}$ erg cm^{-2} s^{-1}).

For a given point in the $\Delta \log L_X - \Delta z$ space, the soft band flux, required to obtain the sky coverage, is derived from the intrinsic hard band luminosity assuming a simple power-law with $\Gamma = 1.8$ for the U subsample, and accounting also for an absorption due to a column density of $N_H = 10^{23.5}$ cm^{-2} at that redshift for the A subsample.

Finally, we computed the total HXLF as the sum of the HXLF of the two subsamples:

$$\phi^{\rm TOT} = \phi^{\rm U} + \phi^{\rm A} \tag{3.4}$$

Errors are computed as the 1σ Poissonian uncertainties (Gehrels 1986) for $\phi^{\rm U}$ and $\phi^{\rm A}$, and then propagated to obtain the errors on $\phi^{\rm TOT}$.

We computed the binned HXLF in 12 bin of the $\Delta \log L_X - \Delta z$ plane (Fig. 3.10 and Tab. 3.4) using Eq. 3.4. Bins were chosen to have an acceptable statistics in each of them. The resulting HXLF is shown in Fig. 3.13. At this stage, we did not include any correction for the redshift incompleteness.

Table 3.4: Number of objects in each $\Delta\log L_X - \Delta z$ bin. Number of obscured objects are between brackets.

ΔLogL$_X$	**z=3-3.19**	**3.19-3.47**	**3.47-3.9**	**3.9-4.3**	**4.3-5.1**
42.85-44	6 (1)	5 (1)	7 (3)	—	—
43-44.5	—	—	—	9 (3)	—
44-44.5	20 (2)	19 (2)	18 (3)	—	—
44-45.1	—	—	—	—	9 (4)
44.5-45.1	14 (8)	13 (5)	14 (2)	—	—
44.5-45.3	—	—	—	7 (2)	—

3.4 Evolutionary models

In order to derive an analytical representation of the HXLF, we assumed different models, starting with a non-evolving, smoothed double power-law:

$$\phi = \frac{\mathrm{d}\Phi(L_X)}{\mathrm{dLog}L} = A\Big[\Big(\frac{L_X}{L_*}\Big)^{\gamma_1} + \Big(\frac{L_X}{L_*}\Big)^{\gamma_2}\Big]^{-1} \tag{3.5}$$

where A is a normalization factor, γ_1 and γ_2 are the slopes of the faint end bright end, respectively, and L_* is the break luminosity (Miyaji et al. 2000; Hasinger et al. 2005).

We then introduced an evolution with redshift. First, we included a pure luminosity evolution (PLE) factor p_{lum}, which preserves the shape of ϕ but shifts it on the luminosity axis (see Fig. 3.12, left panel). The analytical expression is the same as Eq. 3.5 but L_* is multiplied by

$$e_{lum}(z) = (\frac{1+z}{1+z_{min}})^{p_{lum}} \tag{3.6}$$

where p_{lum} is the luminosity evolution factor and $z_{min} = 3$.

Then, we investigated the case of a pure density evolution (PDE), in which the normalization of the HXLF is a function of redshift (Fig. 3.12, left panel) multiplying Eq. 3.5 by a factor

$$e_{den}(z) = (\frac{1+z}{1+z_{min}})^{p_{den}} \tag{3.7}$$

where p_{den} is the density evolution factor.

We also accounted for an independent luminosity and density evolution (ILDE, Yencho et al. 2009), which acts independently on the luminosity and normalization (Fig. 3.12, central panel). Therefore, both the evolutionary factors in Eq. 3.6 and 3.7 were applied simultaneously on Eq. 3.5.

Aird et al. (2010) proposed a luminosity and density evolution model (LADE) similar to the ILDE model, except for a different parametrization of the density evolution and a more complex luminosity evolution, which is assumed to experience a smooth transition between two different regime at a critical redshift. We fitted the LADE model but have slightly changed the analytical form. Since the best-fitting critical redshift found by Aird et al. (2010, $z_c = 0.75 \pm 0.09$) is far from the redshift range probed by this work, we assumed a luminosity evolution with a single slope, parametrized by Eq. 3.6. As for the density evolution, we multiplied the normalization A by

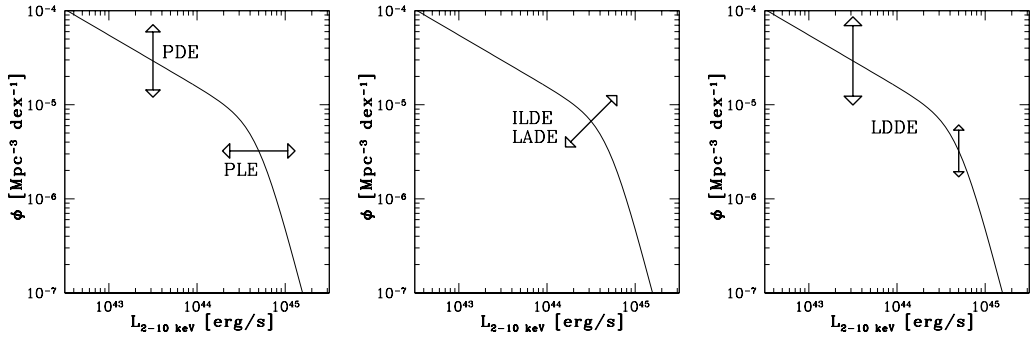

Figure 3.12: Schematic of the HXLF evolution according to the different analytical models described in § 3.4.

$$e_{den}(z) = \frac{10^{p_{den}(1+z)}}{10^{p_{den}(1+z_{min})}}. \tag{3.8}$$

In this form, the parameter A refers to the normalization at $z = 3$, as for the other models.

Finally, we assumed a luminosity-dependent density evolution (LDDE, Schmidt & Green 1983) model. In this case, following the parametrization by Hasinger et al. (2005), Eq. 3.5 is multiplied by

$$e_{den}(z, L) = (\frac{1+z}{1+z_{min}})^{p_{den}+\beta(\ \mathrm{logL}-44)} \tag{3.9}$$

where β is the parameter that accounts for the luminosity dependency. In this case, unlike the previous ones, the shape of the HXLF changes with redshift (Fig. 3.12, right panel).

3.5 Fitting procedure and results

We evaluated the best-fit parameters for the different models using the unbinned maximum-likelihood method (Marshall et al. 1983). The best-fitting parameters for each analytical model $\phi(z, L_X)$ are those which minimize the expression

$$\mathcal{L} = -2\sum_{i=1}^{N} \ln[\phi(z_i, L_i)] + 2\iint \phi(z, L)\Omega\Theta\frac{\mathrm{d}V}{\mathrm{d}z}\ \mathrm{d}z\ \mathrm{d}L \tag{3.10}$$

where N is the total number of sources, and the double integral is computed over the entire $\Delta\log L_X - \Delta z$ region ($3 < z \leq 5.1$ and $42.85 < \mathrm{Log}L_X < 45.3$). Minimization and error analysis were performed with the MINUIT library [3]. Confidence regions are computed at the 1σ c.l. by varying each parameter around the best-fitting value, leaving all the others parameters free, until $\Delta\mathcal{L} = 1$.

Since the sky coverage is different for the U and A subsamples, we minimized $\mathcal{L}^{TOT} = \mathcal{L}^U + \mathcal{L}^A$, where Ω^U and Ω^A are used for the respective subsamples. This is equivalent to perform a simultaneous fit to the two subsamples.

[3] http://lcgapp.cern.ch/project/cls/work-packages/mathlibs/minuit/index.html

Table 3.5: Best-fit parameters.

MODEL	A	L_*	γ_1	γ_2	p_{lum}	p_{den}	β	2DKS
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
PLE	$0.65^{+0.06}_{-0.06}$	$6.56^{+2.38}_{-2.22}$	$0.21^{+0.16}_{-0.20}$	$2.58^{+0.75}_{-0.60}$	$-3.73^{+0.77}_{-0.92}$	—	—	0.05
PDE	$1.10^{+0.11}_{-0.11}$	$5.26^{+1.06}_{-1.20}$	$0.22^{+0.13}_{-0.16}$	$3.79^{+1.08}_{-0.87}$	—	$-6.00^{+0.84}_{-0.87}$	—	0.38
ILDE	$1.13^{+0.11}_{-0.11}$	$5.13^{+1.32}_{-1.53}$	$0.21^{+0.13}_{-0.17}$	$3.75^{+1.10}_{-0.91}$	$0.13^{+0.94}_{-0.81}$	$-6.13^{+1.17}_{-1.23}$	—	0.38
LADE	$1.08^{+0.11}_{-0.11}$	$5.10^{+1.33}_{-1.54}$	$0.21^{+0.13}_{-0.18}$	$3.74^{+1.11}_{-0.91}$	$0.16^{+0.95}_{-0.82}$	$-0.57^{+0.11}_{-0.11}$	—	0.46
LDDE	$1.05^{+0.10}_{-0.10}$	$5.24^{+1.05}_{-1.19}$	$0.28^{+0.16}_{-0.19}$	$3.87^{+1.08}_{-0.88}$	—	$-6.43^{+1.12}_{-1.17}$	$1.18^{+2.06}_{-2.00}$	0.42

(1) model; (2) normalization in units of 10^{-5} Mpc^{-3}; (3) knee luminosity in units of 10^{44} erg s^{-1}; (4) faint and (5) bright end slope; (6) luminosity and (7) density evolutionary factor; (8) luminosity-dependency factor of the density evolution; (9) two dimensional Kolmogorov-Smirnov test probability.

We determined the normalization parameters A^U and A^A by imposing that the total number of sources is the observed one, separately for the U and A subsamples. Their sum gives the normalization A of the total HXLF. Errors on the normalization are computed separately for the two subsamples, accounting for the statistical errors on the number of sources, and then propagated to derive the error on A. Best-fitting models are plotted in Fig. 3.13 and the parameters are reported in Tab. 3.5.

Unlike the χ^2 test, the maximum likelihood method gives no information about the goodness of the fit. We evaluated it by using the bidimensional Kolmogorov-Smirnov test (2DKS; Peacock 1983; Fasano & Franceschini 1987). We mention that a value of the 2DKS probability $\gtrsim 0.2$ may not be accurate, but still means that the data and the model predictions are not significantly different (Press et al. 1992). Also, the formal KS test derivation requires the two compared samples to be independent. This is not the case, since we will compare data and models derived by fitting the data themselves. Therefore, the resulting 2DKS test values are expected to be overestimated. However, following e.g. Miyaji et al. (2000), we chose to use these probabilities as goodness-of-fit indicators, as we are more interested in comparing different models rather than in the absolute probabilities, and not to apply the formally-correct, but computationally very expensive, treatment, which would involve large sets of Monte-Carlo simulations.

3.5.1 Introducing the completeness factor

So far we have not considered the redshift incompleteness of the sample and computed the HXLF assuming $\Theta = 1$ in Eq. 3.3. We then searched for a reasonable method to estimate it. A procedure adopted by some authors (e.g. Barger & Cowie 2005; Yencho et al. 2009) to provide upper limits to the complete HXLF is to count all sources with no redshift information in each $\Delta \log L_X - \Delta z$ bin. However, this method would result in very loose constraints on the HXLF. Other authors developed procedures based on optical luminosity (e.g. Matute et al. 2006).

There are 31, 28 and 6 soft-band detected sources in the CDFS, C-COSMOS and XMM-COSMOS surveys which do not have a spectroscopic or photometric redshift. Hiroi et al. (2012) reported that for 6 AGN candidates in the SXDS neither a spectroscopic nor a photometric redshift was available. Since the catalogue of the optical survey in this region is not public, we do not know which are those 6 X-ray sources and we exclude them from the computation of the completeness correction.

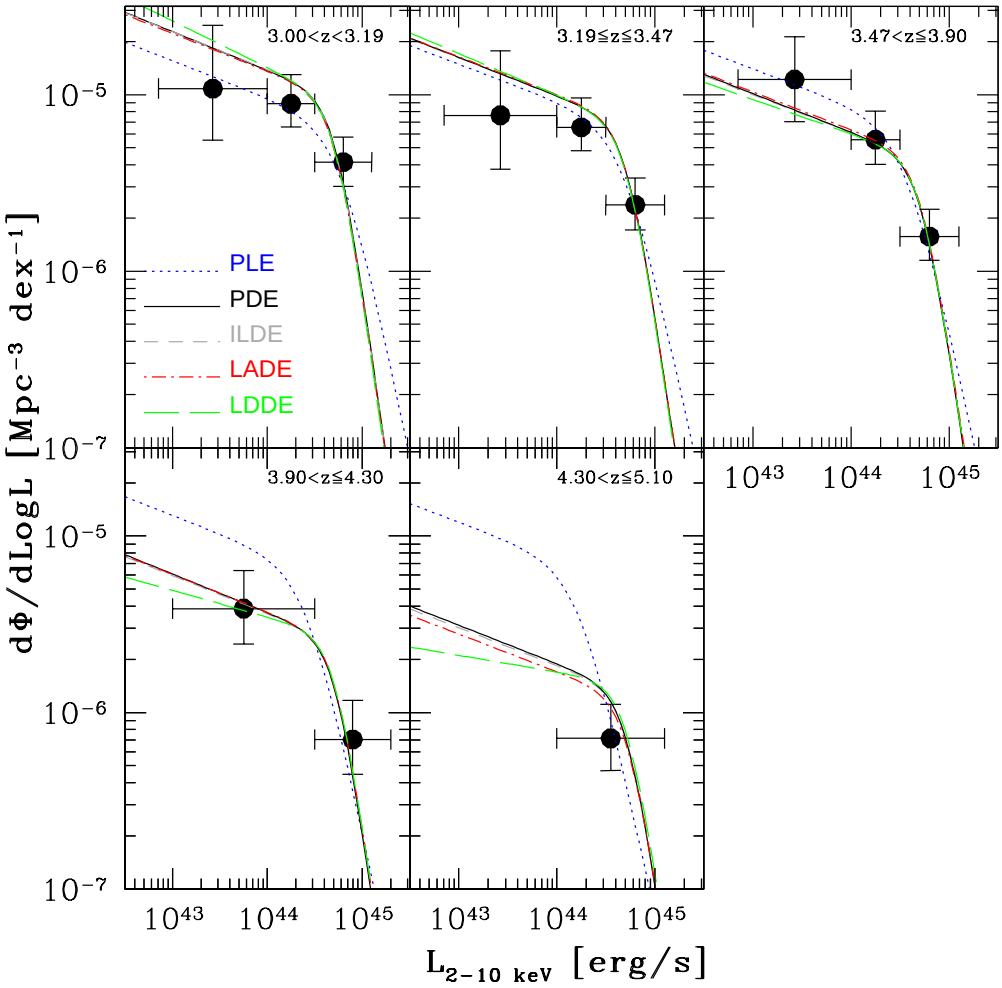

Figure 3.13: Binned luminosity function. The best fit for the models considered in § 3.4 are evaluated at the mean redshift of each bin and plotted with different line styles and colours. The best-fitting ILDE model (short-dashed line) is barely visible over the PDE model (solid line), having very similar best-fitting parameters (see Tab. 3.5 and text). No correction for redshift incompleteness has been included.

We do not expect significant changes if they would be included, since they represent a tiny fraction of the overall sample.

We made the extreme assumption that all the 65 sources are at $3 < z < 5.1$ and applied the following method to derive $\Theta(z, L_X, N_H)$, i.e. the fraction of sources in the high-redshift sample with redshift information in a given $\Delta\log L_X - \Delta z$ bin: each source with no redshift information, but for which the flux F is known, is counted simultaneously in the A and U subsample with a weight ($w_A(F)$ and $w_U(F) = 1 - w_A(F)$, respectively) corresponding to the fraction of absorbed ($w_A(F)$) or unabsorbed ($w_U(F)$) sources with redshift at similar fluxes. Each weighted source is then spread in redshift accordingly to the normalized redshift distribution of the sources with redshift in the relative subsample ($P_A(z)$ and $P_U(z)$). The redshift distribution were evaluated in the same bins used for the binned HXLF. Different choices of bins do not result in significant change of the results and have the same level of arbitrariness. The luminosity was then computed for all the weighted sources in the two subsamples in each redshift bin, at the median redshift of the bin. Finally, knowing the number of sources with redshift and having estimated the number of those without in each $\Delta\log L_X - \Delta z$ bin, we derived $\Theta(z, L_X, N_H)$. We note that using this procedure all the 65 sources are included, since

$$\sum_{i=1}^{65} \sum_{z} (w_A(F_i) P_A(z) + w_U(F_i) P_U(z)) = 65 \tag{3.11}$$

The effect of this correction is shown in Fig. 3.14, where the HXLF is also plotted separately for the two absorption-based subsamples (upper panels). As expected, incompleteness is more severe at low X-ray luminosity, usually associated to optical faintness. Results from Brusa et al. (2009, XMM-COSMOS, green square) Civano et al. (2011, *Chandra*-COSMOS, green triangle) and Fiore et al. (2012, 4 Ms CDF-S, green stars) in the 4 Ms CDF-S, are also shown for comparison. In particular, Fiore et al. (2012) selected a sample of high redshift objects using spectroscopic and photometric redshifts, and a colour selection. Vito et al. (2013) applied a different selection, based on X-ray detection only, and a careful verification of the redshifts on the same field, which resulted in the discrepancy at the faint end of the luminosity function, dominated by the CDFS sample, when no completeness correction is applied. After the inclusion of completeness, our estimate gets closer to the Fiore et al. (2012). Models from Hasinger et al. (2005, LDDE), La Franca et al. (2005, LDDE), Aird et al. (2010, LADE) and Ueda et al. (2014, LDDE) are also plotted. Before the completeness correction, the faint end of the HXLF shows a flattening similar to that of the Hasinger et al. (2005) model, but the corrected HXLF shows a much steeper slope.

We computed the binned HXLF and repeated the fitting procedure described in § 3.5 including the effect of incompleteness (Fig. 3.15). Best-fitting parameters are reported in Tab. 3.6.

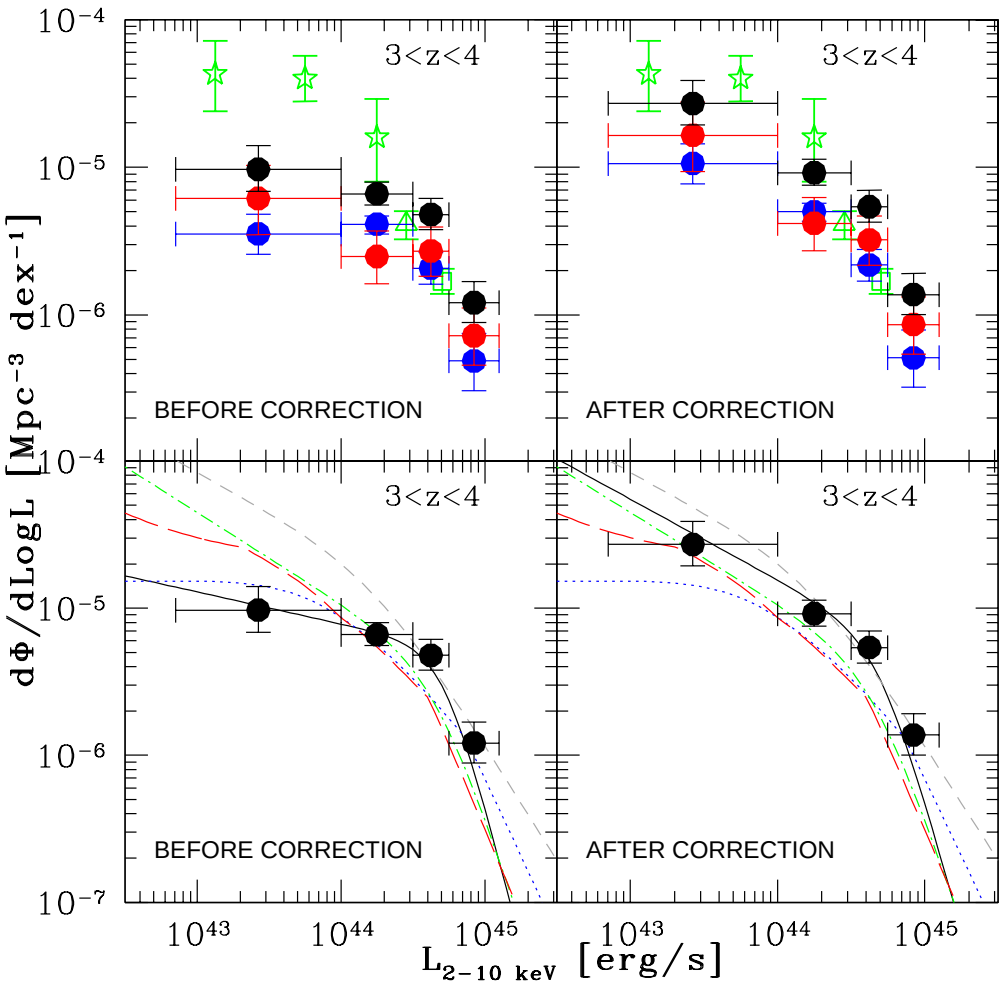

Figure 3.14: HXLF in a representative redshift bin ($3 < z < 4$) computed before (*left panels*) and after (*right panels*) the correction for incompleteness. In the upper panels, the HXLF of U (blue circles) and A (red circles) subsamples are plotted to show the differential correction. The total HXLF (black circles) is computed following Eq. 3.4. Results from Brusa et al. (2009, XMM-COSMOS, green square) Civano et al. (2011, *Chandra*-COSMOS, green triangle) and Fiore et al. (2012, 4 Ms CDF-S, green stars) in the 4 Ms CDF-S, are also shown for comparison. In the lower panels, we compare the HXLF with results from literature. The LDDE model by Hasinger et al. (2005), where obscured (Compton-thin) AGN and an exponential decline in the space density are accounted following Gilli et al. (2007), is shifted to the hard band and plotted as a blue dotted line. We also show models by La Franca et al. (2005, LDDE, grey short-dashed line), Aird et al. (2010, LADE, green dot-dashed line) and Ueda et al. (2014, LDDE, red long-dashed line) and added our best-fitting model (PDE, black solid lines) before and after the correction.

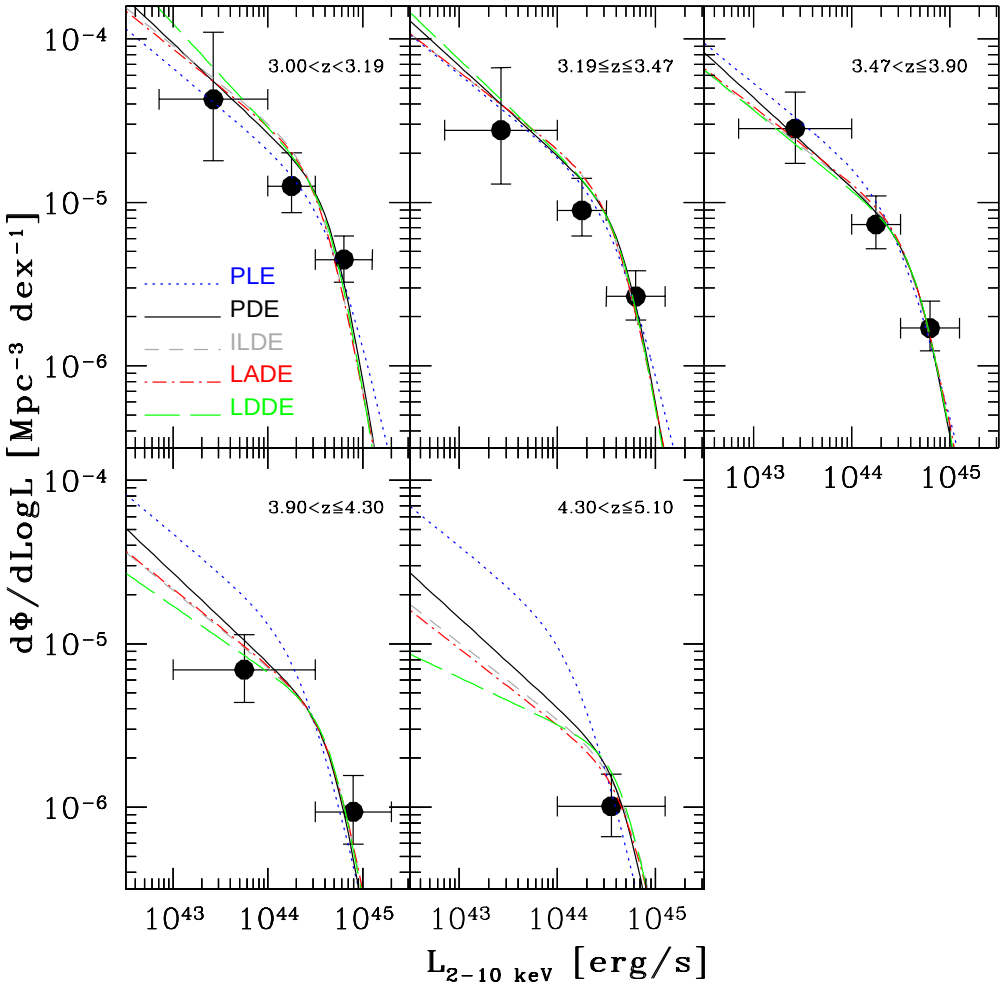

Figure 3.15: Same as Fig. 3.13 but including the correction for redshift incompleteness (see § 3.5.1).

Table 3.6: Best-fit parameters to the HXLF, including the correction for redshift incompleteness.

MODEL (1)	A (2)	L_* (3)	γ_1 (4)	γ_2 (5)	p_{lum} (6)	p_{den} (7)	β (8)	2DKS (9)
PLE	$1.00^{+0.10}_{-0.10}$	$5.22^{+2.91}_{-2.43}$	$0.49^{+0.19}_{-0.28}$	$2.55^{+0.79}_{-0.55}$	$-3.19^{+0.61}_{-0.67}$	—	—	0.20
PDE	$1.25^{+0.13}_{-0.13}$	$5.03^{+1.37}_{-1.55}$	$0.55^{+0.14}_{-0.19}$	$3.62^{+1.13}_{-0.86}$	—	$-5.68^{+0.87}_{-0.90}$	—	0.44
ILDE	$2.00^{+0.21}_{-0.21}$	$3.63^{+2.11}_{-1.61}$	$0.47^{+0.25}_{-0.22}$	$3.15^{+1.22}_{-0.70}$	$1.19^{+1.52}_{-1.29}$	$-7.19^{+1.88}_{-2.06}$	—	0.20
LADE	$1.84^{+0.19}_{-0.19}$	$3.68^{+2.08}_{-1.61}$	$0.47^{+0.19}_{-0.25}$	$3.17^{+1.22}_{-0.71}$	$1.12^{+1.45}_{-1.25}$	$-0.65^{+0.17}_{-0.18}$	—	0.23
LDDE	$1.19^{+0.11}_{-0.11}$	$4.92^{+1.37}_{-1.53}$	$0.66^{+0.17}_{-0.21}$	$3.71^{+1.12}_{-0.84}$	—	$-6.65^{+1.28}_{-1.32}$	$2.40^{+2.33}_{-2.31}$	0.27

(1) model; (2) normalization in units of 10^{-5} Mpc^{-3}; (3) knee luminosity in units of 10^{44} erg s^{-1}; (4) faint and (5) bright end slope; (6) luminosity and (7) density evolutionary factor; (8) luminosity-dependency factor of the density evolution; (9) two dimensional Kolmogorov-Smirnov test probability.

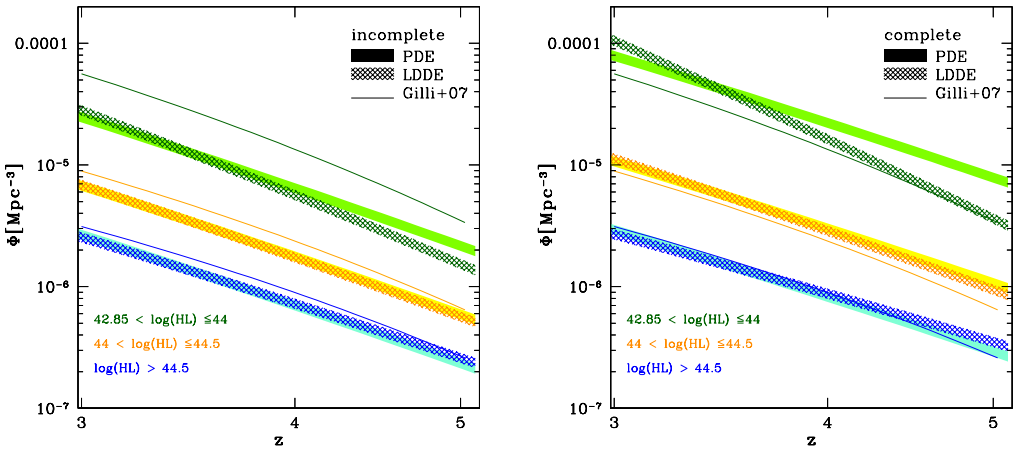

Figure 3.16: Space density in three different, colour-coded luminosity bins, estimated from the best-fitting PDE (filled, lighter stripes) and LDDE (gridded, darker stripes) models. In the left panel, we assumed $\Theta = 1$ (i.e. no redshift completeness correction), while in the right panel we applied the correction for redshift incompleteness (as in § 3.5.1). The three solid lines are the predictions from the Gilli et al. (2007) X-ray background synthesis model computed including also the most Compton-thick AGN ($N_H > 10^{25}$ cm^{-2}).

3.6 Expected space density from HXLF best-fitting models

Given the best-fitting HXLF evolutionary models $\phi(\log L_X, z)$ in § 3.5, we can computed the expected space density $\Phi^{exp}(z)$ in a given luminosity bin $\Delta \log L_X$ for each representation as

$$\Phi^{exp}(z) = \int_{\Delta \log L_X} \phi \; d\log L_X \tag{3.12}$$

The space density expected from the best-fitting PDE and LDDE models is plotted as a function of redshift in Fig. 3.16 in three different luminosity bins, before (left panel) and after (right panel) correcting for the redshift incompleteness (see § 3.5.1). The width of the stripes accounts only for the uncertainties on the normalization of each model (i.e. on the number of sources).

We also show the expectations from the Gilli et al. (2007) X-ray background synthesis model, including also the most Compton-thick AGN ($N_H > 10^{25}$ cm^{-2}) and adopting a high-redshift decline of the AGN space density[4].

[4] http://www.bo.astro.it/∼gilli/xvol.html

3.7 Binned estimates of the space density

A binned representation of the space density Φ of an extragalactic population of objects can be derived using the $1/V_{max}$ method (Schmidt 1968; Avni & Bahcall 1980):

$$\Phi = \frac{\mathrm{d}N}{\mathrm{d}V} = \sum_{i=1}^{N} \frac{1}{V_{max,i}} = \sum_{i=1}^{N} \frac{1}{\int \Omega \Theta \frac{\mathrm{d}V}{\mathrm{d}z}\ \mathrm{d}z} \tag{3.13}$$

where N is the number of objects in the $\Delta \mathrm{Log}L - \Delta z$ bin of interest and $\Theta = \Theta(z, L_{x,i}, N_H)$ is the completeness factor, as defined in § 3.5.1.

We computed the binned space density of $\mathrm{logL_X} > 44.15$ AGN as $\Phi = \Phi_U + \Phi_A$, where Ω^{U} and Ω^{A} are used to compute Φ_U and Φ_A, respectively. Similarly to Hiroi et al. (2012), errors are estimated as

$$\delta\Phi_U = \sqrt{\sum_{i=1}^{N_U} \frac{1}{V_{max,i}^2}} \tag{3.14}$$

where N_U is the number of sources in the U subsample, for Φ_U and likewise for Φ_A. Then, they are propagated to obtain the error on Φ. The results are plotted in Fig. 3.17 and discussed in § 3.9.

In principle, we could also use the Page & Carrera (2000) method to derive a binned HXLF in a given luminosity bin and then multiplying it by the corresponding ΔLogL to obtain Φ. However, we computed the space density in a luminosity bin ($\mathrm{logL_X} > 44.15$) even larger that those used in § 3.3 and the assumption that ϕ is not varying in that bin would strongly affect the evaluation of the binned points. Instead, since the fluxes corresponding to such luminosities are expected to be higher than the flux limit of the surveys, we can safely use the $1/V_{max}$ method (see § 3.3).

3.8 Number counts

In order to check the conclusions on the number counts derived in § 2.3.4 using the CDFS sample only, we computed the logN-logS for the entire sample of 141 $z > 3$ AGN. The same procedure adopted in § 2.3.4 was applied, assuming either Ω^{U}, if one source belongs to the U subsample, or Ω^{A} as the proper sky coverage. The logN-logS distribution includes a number of AGN slightly smaller than those used by Kalfountzou et al. (2014) for their number counts, but, having included data from the 4 Ms CDFS, we can probe fluxes a factor of ~ 10 fainter.

The logN-logS of our high-redshift sample is shown in Fig. 3.18. At $z > 3$ we confirm the agreement with the Gilli et al. (2007) model including the decline of the AGN space density, as found in § 2.3.4. Using only the 4 Ms CDFS sample, at $z > 3$ the number counts were slightly above the Gilli et al. (2007) model. The new logN-logS with the improved statistics is now in good agreement with that model at $z > 4$ as well.

In Fig. 3.19 and 3.20 we report the number counts at $z > 3$ and $z > 4$, separately for the unabsorbed and absorbed subsamples and the predictions from the models

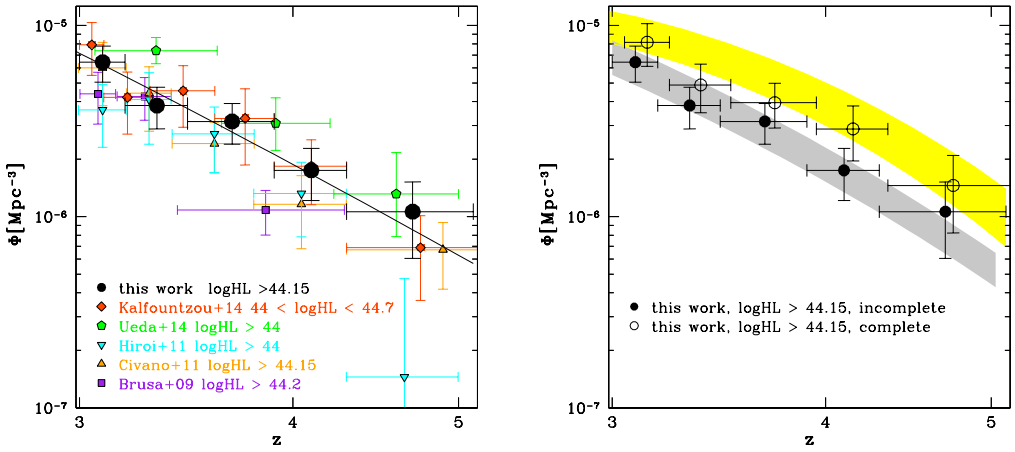

Figure 3.17: Left panel: our estimate of the AGN space density (LogL > 44.15, not corrected for incompleteness; black circles) as a function of redshift is compared with results from Kalfountzou et al. (2014, 44 < LogL < 44.7, red diamonds), Ueda et al. (2014, LogL > 44, green pentagons), Hiroi et al. (2012, LogL > 44, cyan downward-pointing triangles), Civano et al. (2011, LogL > 44.15, orange upward-pointing triangles) and Brusa et al. (2009, LogL > 44.2, purple squares). The solid line is the best-fitting PDE model. Right panel: AGN space density before (filled circles) and after (empty circles) the correction for redshift incompleteness. The grey and yellow stripes are the expectations from the X-ray background synthesis models by Gilli et al. (2007) and Shi et al. (2013), respectively, computed at LogL > 44.15 between the two extreme cases of accounting only for Compton-thin ($N_H < 10^{24}$ cm^{-2}; lower bounds) and including also Compton-thick (upper bounds) AGN.

by Gilli et al. (2007), Treister et al. (2009)[5] and Akylas et al. (2012)[6]. By comparing our points with the Gilli et al. (2007) predictions, we found the logN-logS of the unabsorbed subsample to be in good agreement with the model (assuming the decline scenario), while the logN-logS of absorbed sources is underestimated by a factor $\sim 1.5-2$ by the model. Since Gilli et al. (2007) assume a constant obscured AGN fraction with redshift and a varying fraction with luminosity, the discrepancy can be probably due to an evolution of the obscured AGN fraction with redshift, a different dependency on the luminosity or a combination of these two effects. This would be in qualitative agreement with the results on the obscured AGN fraction which will be presented in § 4. However, the discrepancy would be sensibly reduced also by considering the errors on the parameters used by Gilli et al. (2007) for their model.

3.9 Discussion and conclusions

Before the correction for redshift incompleteness, the 2DKS test returns similar high values for the PDE, ILDE, LADE and LDDE models (Tab. 3.5). However, the common parameters are very similar and, since the ILDE, LADE and LDDE models have an additional free parameter consistent with zero (see Tab. 3.5), we conclude that these three models mimic the behaviour of the PDE model. Indeed, the PDE, ILDE and LADE models plotted in Fig. 3.13 are almost completely overlapping. As for the LDDE model, we note that the evolution dependency on the luminosity is more effective in changing the shape of the HXLF at low luminosities. Therefore, we ascribe the weak constraints we found on the β parameter to the poor sampling of the low-luminosity regime.

After the correction for redshift incompleteness, the PDE model has the highest 2DKS test value. The LDDE model is still consistent with the data. The best-fitting ILDE and LADE models are again consistent with no luminosity evolution, although at a lower significance level than in § 3.5 (see Tab. 3.6). According to the 2DKS probability, these models, as well as the LDDE and PLE ones, are acceptable with a much lower significance than in the previous case..

We conclude that the evolution of the $z > 3$ HXLF is dominated by a negative density term. A luminosity evolution term which is present in the other parameterization of the luminosity function is either inconsistent (i.e. PLE) or not required (ILDE, LADE and LDDE) by the data. A better sampling of the high-redshift HXLF faint end ($\log L_X < 44$), where the effect of a LDDE is more noticeable, is required to assess the evolution of the $z > 3$ AGN population. Recently, Ueda et al. (2014) presented the X-ray luminosity function of AGN from $z = 0$ to 5 and chose a LDDE model to analytically describe the evolution, fixing the model parameters at high redshift ($z > 3$). Instead, we focused on the particular redshift range $3 < z < 5.1$. Therefore the best-fitting parameters are not weighted by lower redshift data, where much larger samples of AGN are available. We also have not

[5]Treister et al. (2009) used a combination of γ-ray, X-ray and mid-IR data to constrain the space density of Compton-Thick AGN.

[6]Akylas et al. (2012) constrained the number of Compton-Thick AGN by fitting the XRB using library of X-ray spectra. The ratio between absorbed and unabsorbed AGN is assumed to decrease from ~ 0.8 at low luminosity to ~ 0.4 at high-luminosity and to not depend on luminosity.

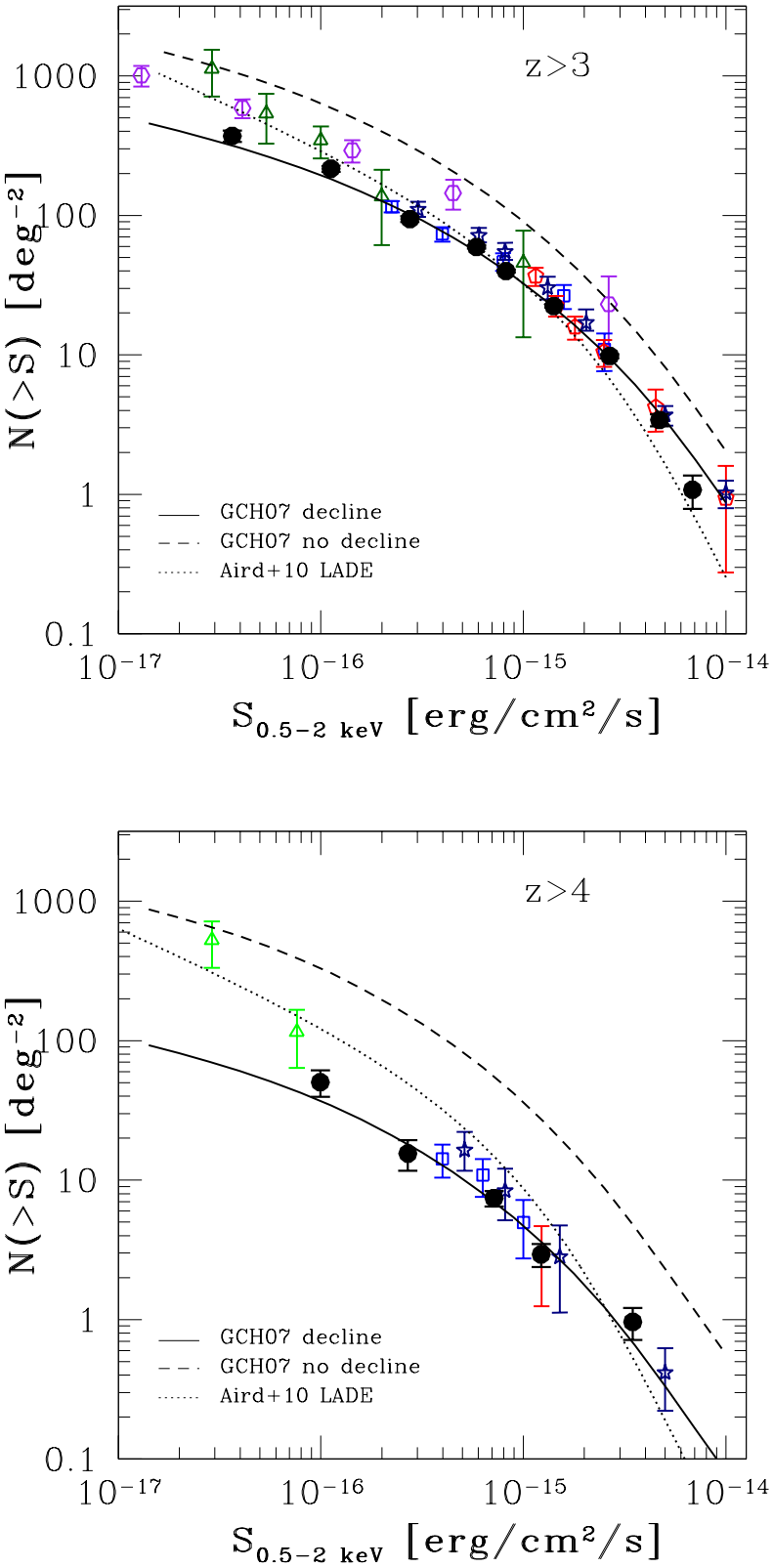

Figure 3.18: Binned LogN-logS of the high-redshift sample, considering all sources ($z > 3$, *left panel*) and only those at $z > 4$ (*right panel*), with associated Poissonian errors. Symbols and lines are the same as in Fig. 2.12. No correction for redshift incompleteness was applied.

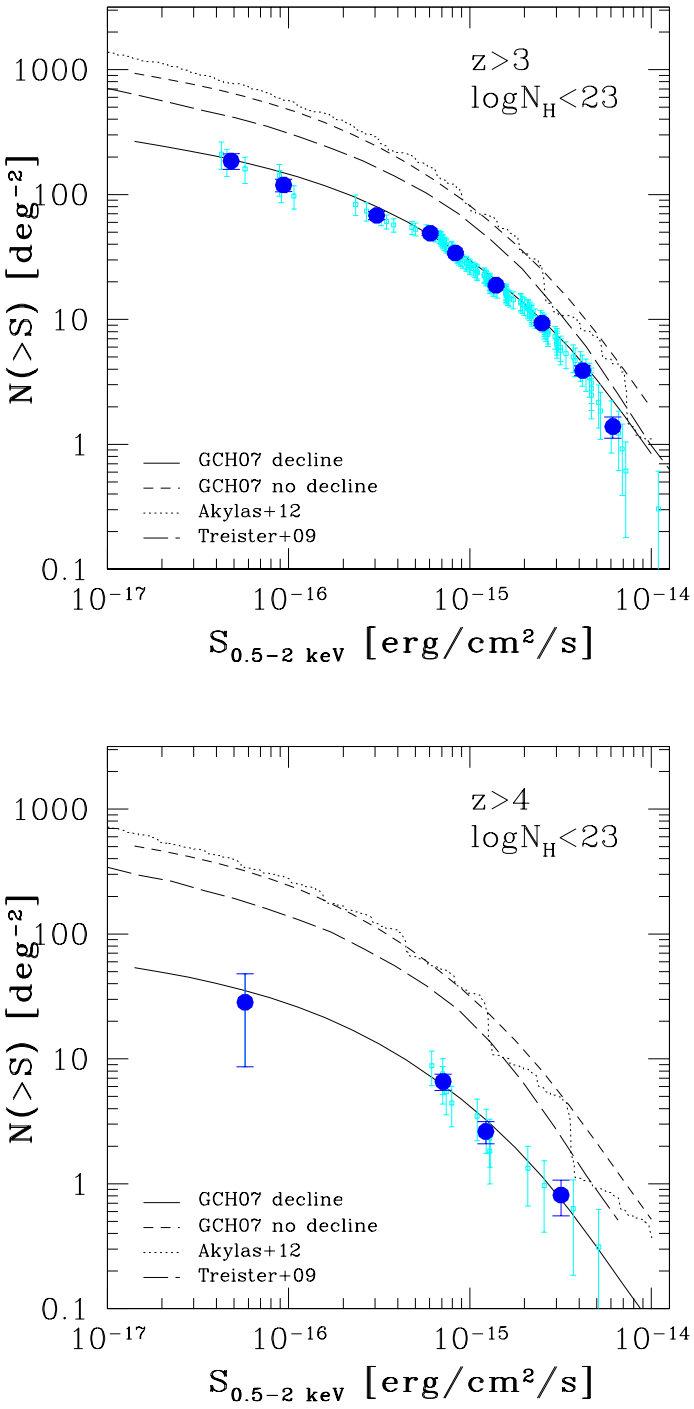

Figure 3.19: Binned LogN-logS (blue filled circles) of the unabsorbed subsample overplotted to the individual points (cyan open squares), at $z > 3$ (*left panel*) and $z > 4$ (*right panel*). No correction for redshift incompleteness was applied. Lines are the predictions from the X-ray background synthesis models reported in the panels.

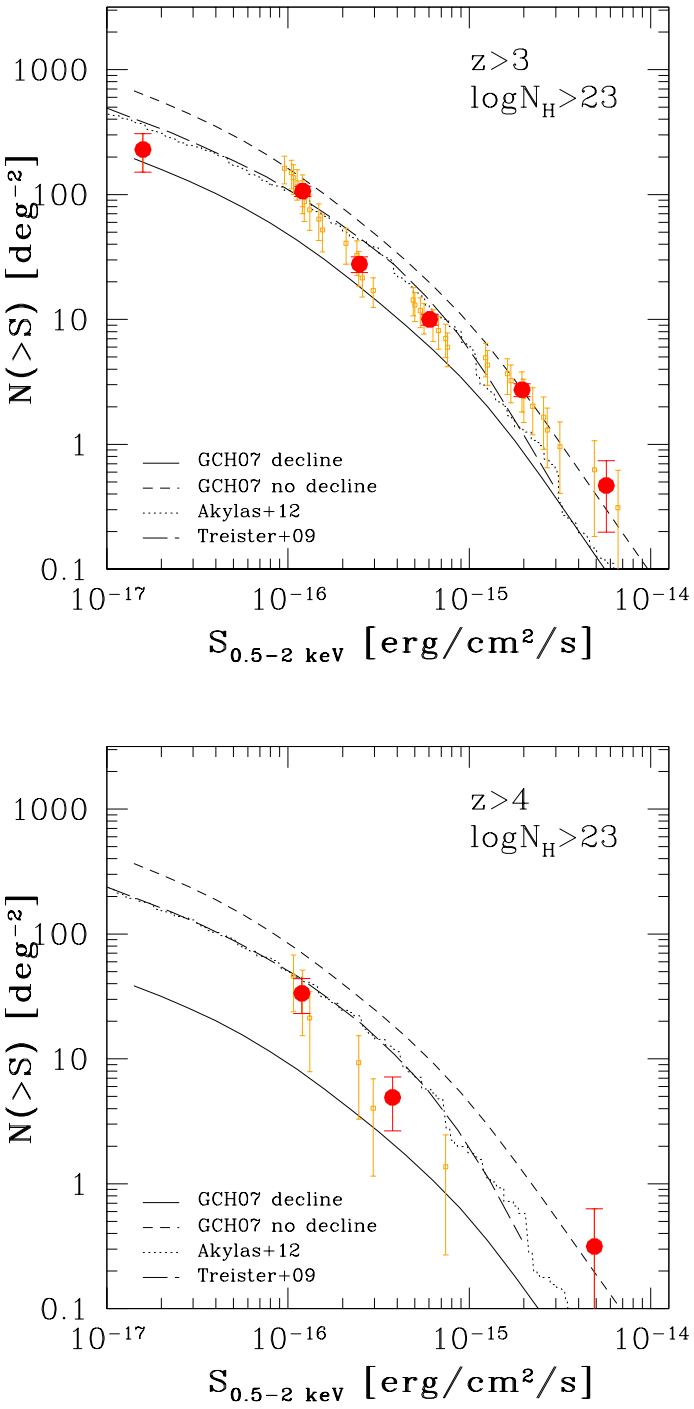

Figure 3.20: Binned LogN-logS (red filled circles) of the absorbed subsample overplotted to the individual points (orange open squares), at $z > 3$ (*left panel*) and $z > 4$ (*right panel*). No correction for redshift incompleteness was applied. Lines are the predictions from the X-ray background synthesis models reported in the panels.

fixed any model parameter.

Ueda et al. (2014) argued that, at $z > 3$, X-ray selected AGN experience an "up-sizing" evolution (i.e. the space density of low luminosity AGN declines with increasing redshift at a slower rate than for high luminosity objects). There is no evidence of such an effect in our sample. On the contrary, there may be hints of an opposite trend (see Fig. 3.16, dashed lines). The origin of the discrepancy could be due to the different sample selection at high-z and/or to the details of the choice of the parameters which are fixed in the Ueda et al. (2014) fits.

In Fig. 3.17 we compare our results at $\mathrm{LogL_X} > 44.15$, not corrected for redshift incompleteness, with those from Kalfountzou et al. (2014, $44 < \mathrm{LogL_X} < 44.7$), Ueda et al. (2014, $\mathrm{LogL_X} > 44$), Hiroi et al. (2012, $\mathrm{LogL_X} > 44$), Civano et al. (2011, $\mathrm{LogL_X} > 44.15$) and Brusa et al. (2009, $\mathrm{LogL_X} > 44.2$), and with the predictions from the X-ray background synthesis models by Gilli et al. (2007) and Shi et al. (2013)[7]. In particular, Hiroi et al. (2012) jointly fitted their and the Civano et al. (2011) points with a powerlaw in the form $\Phi \propto (1+z)^p$, finding a decline in the space density of high-redshift AGN with a slope of $p = -6.2 \pm 0.9$. Our best-fitting PDE model returns a very similar evolution (see Tab. 3.5 and 3.6), although a larger dataset (which includes the Hiroi et al. sample) was used and a different method (i.e. maximum likelihood fit to unbinned data in the $z - \mathrm{logL_X}$ space) was adopted. Another hint in favour for a declining scenario comes from the logN-logS of the high-redshift sample, which is in agreement with the model by Gilli et al. (2007) if the decline of the AGN space density at high redshift is assumed.

In conclusion, we found that the space density of luminous AGN decreases by a factor of ~ 10 from $z = 3$ to 5. Larger samples of AGN, especially of low luminosity ($\mathrm{logL_X} < 44$) and at $z > 4$, are necessary to constrain the space density of high-redshift AGN and will be provided by future surveys like the additional *Chandra* 3 Ms observation in the CDFS (PI: N. Brandt) and the *Chandra*-COSMOS Legacy survey (PI: F. Civano).

[7]The model by Shi et al. (2013) was built by fitting number counts, redshift distributions and local luminosity functions in the X-ray and IR bands with 19 free parameters, finding the obscured AGN fraction to increase with redshift and decrease with luminosity.

Chapter 4

Obscured AGN fraction at high redshift

The evolution over cosmic time of the obscured AGN fraction is one of the main open questions in our knowledge of AGN evolution (see § 1.3 and § 1.4). In particular, the value of this quantity at high redshift is not known with sufficient accuracy. In this chapter, we use the sample of high-redshift AGN described in § 3.2 to estimate the obscured AGN fraction at $3 < z < 5$ (§ 4.1) and compare it with previous results at low and high redshift (§ 4.2). These results are presented in Vito et al (submitted).

4.1 Method

Having computed in § 3.3.1 the binned HXLF separately for the U (ϕ^U) and A (ϕ^A) subsamples, we can estimate the obscured AGN fraction (F_{23}) in a $\Delta\log L_X - \Delta z$ bin as

$$F_{23} = \frac{\phi^A}{\phi^A + \phi^U} \tag{4.1}$$

where the subscript is added to stress that the separation between the two subsample is at $\log N_H = 23$. However, by using the same $\Delta\log L_X - \Delta z$ bins as in § 3.3, the errors on F_{23} resulted too large to draw reasonable conclusions, because of the low number of sources (especially belonging to the A subsample) in such narrow bins. We chose to compute F_{23} over the entire redshift range $3 < z < 5.1$ in four luminosity bins (Fig. 4.1).

Since the Page & Carrera (2000) method assumes that the luminosity function does not vary in the considered bins, as discussed in § 3.3, the result we present in this chapter should be considered as a rough estimate of the average obscured AGN fraction in that redshift range. In § 3.7 we derived the space density of high-redshift AGN using the $1/V_{max}$ method. The proper way to estimate F_{23} would be to use the space density itself instead of ϕ in Eq. 4.1. However, obscured sources are on average closer to the flux limit of a survey than unobscured AGN at a given luminosity. The $1/V_{max}$-related bias reported in § 3.3 is therefore more effective for the A subsample than for the U subsample, and this fact would strongly affect the evaluation of the obscured AGN fraction, especially at low luminosities. All values

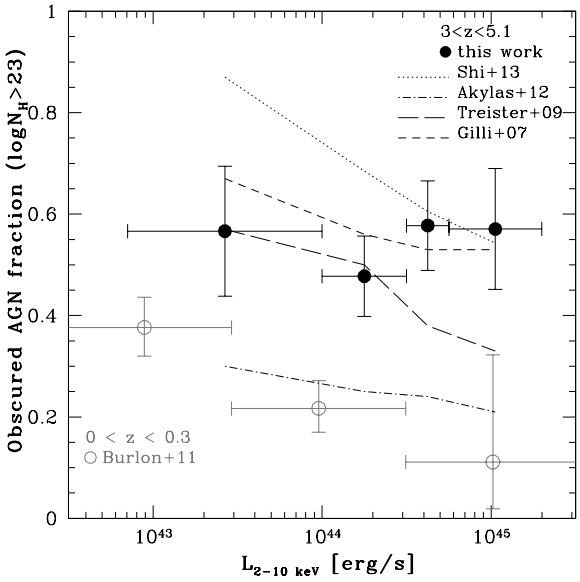

Figure 4.1: Obscured AGN fraction as a function of luminosity. The dotted, dot-dashed, long-dashed and short-dashed lines are the predictions from the X-ray background synthesis models by Shi et al. (2013), Akylas et al. (2012), Treister et al. (2009) and Gilli et al. (2007), respectively, computed at the median of the considered redshift range.

of F_{23} were obtained after the correction for redshift incompleteness (see § 3.5.1). We checked that no significant difference would be obtained if we had not applied that correction.

4.2 Results and discussion

The obscured AGN fraction at $3 \leq z \lesssim 5$ seems to be constant with luminosity: a simple χ^2 fit to the filled circles in Fig. 4.1 returns a value of $F_{23} = 0.54 \pm 0.05$. This is in contrast to the decreasing obscured fraction with luminosity reported in many works, mainly at lower redshift (see § 1.3.1). However, heavily-obscured ($\log N_H \gtrsim 23.5$), low luminosity ($\log L_X \lesssim 44$) AGN at $z > 3$ cannot be currently detected even by the deepest X-ray surveys. Indeed, in our sample there is a clear deficiency of such objects with respect to more luminous ($\log L_X \gtrsim 44$) AGN affected by a similar level of obscuration, as shown in Fig. 2.5 (right panel). The obscured AGN fraction could then be larger than our estimates, especially in the first luminosity bin of Fig. 4.1.

We also compared our points with the predictions from the Gilli et al. (2007), Treister et al. (2009)[1], Akylas et al. (2012)[2] and Shi et al. (2013)[3] X-ray background synthesis models (see Fig. 4.1). At low luminosities, our points are consistent with

[1] http://agn.astroudec.cl/j_agn/main

[2] http://indra.astro.noa.gr/xrb.html

[3] http://5muses.ipac.caltech.edu/5muses/EBL_model/num_den_red.html

the Treister et al. (2009) and Gilli et al. (2007) models, while at high luminosities we find a good agreement with the Gilli et al. (2007) and Shi et al. (2013) predictions.

When comparing our points (filled circles in Fig. 4.1) with results in the local Universe from Burlon et al. (2011), who studied a complete sample of AGN detected by *Swift*–BAT, we found a positive evolution of the obscured AGN fraction from $z = 0$ to $z > 3$, which is stronger at high luminosities ($\log L_X > 44$), even considering that the low luminosity bin at $z > 3$ could be affected by some incompleteness. This result agrees with the increasing fraction of absorbed AGN with redshift reported in literature (e.g. La Franca et al. 2005; Akylas et al. 2006; Treister & Urry 2006; Hasinger 2008; Burlon et al. 2011; Ueda et al. 2014). This finding fits well in the wider scenario described in § 1.4.1, in which the different triggering mechanism between high- and low-luminosity AGN causes a different evolution of the obscured fraction from low to high redshift.

In order to search for a possible evolution from $z = 3$ to $z = 5$, in Fig. 4.2 we plot the obscured AGN fraction against redshift for two luminosity bins. The large error bars, the narrow redshift range, the relatively low number of (obscured) sources and the caveats reported in § 4.1 prevent us from drawing strong conclusions. We note that an absence of evolution at high redshift would be in agreement with Hasinger (2008), who suggested that the Type-2 AGN fraction saturates at $z \geq 2$ (but see also Gilli et al. 2010). The assess the evolution of obscuration at $z > 3$, larger samples of AGN are needed. The on-going 7 Ms CDFS (PI: W.N. Brandt) and 2.8 Ms *Chandra* COSMOS Legacy (PI: F. Civano) surveys will provide tens and hundreds of new high-redshift sources, respectively. The inclusion of such objects on the one hand will increase the statistics, allowing us to investigate the obscured AGN fraction in bins of redshift at $z > 3$ and probe the presence or absence of an evolution, on the other hand it will populate the low-luinosity regime of the AGN population, so far poorly sampled.

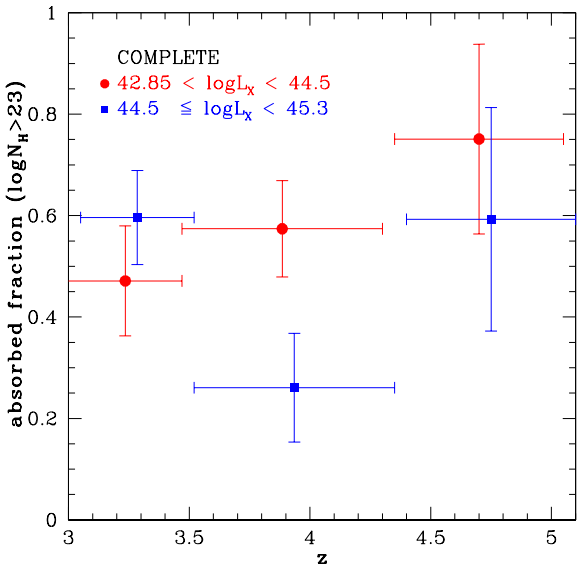

Figure 4.2: Obscured AGN fraction as a function of redshift for two luminsity ranges. We applied the correction for redshift incompleteness, as described in the text.

Chapter 5

The gas content of AGN-hosting galaxies at low redshift ($z < 1$)

In this chapter, we investigate the gas content of a sample of several hundreds AGN host galaxies at $z < 1$ and compared it with a sample of inactive galaxies (star forming and quiescent), matched in bins of stellar mass and redshift. The results may explain the connection between nuclear activity and star formation in the host galaxy (see § 1.5.2).

The molecular gas content can be inferred from the luminosity of the CO millimeter transitions, by assuming a proper CO-to-H_2 conversion factor. However, CO observations are very time consuming, and surveys of large samples of objects are extremely difficult and time demanding. Alternatively, the total (molecular and atomic) gas mass can be derived from the dust content, inferred from the FIR-submm SED, by assuming a dust-to-gas ratio (DGR; e.g. Eales et al. 2010; Leroy et al. 2011; Magdis et al. 2011). The uncertainties on the dust-to-gas ratio and its dependence on metallicity are similar to those affecting the CO-to-H_2 conversion factor, making the two methods comparable in terms of accuracy, at least at metallicities 12+log(O/H)>8.0 (Bolatto et al. 2013; Rémy-Ruyer et al. 2013).

For the first time, we exploit the dust method for measuring the gas masses in AGN host galaxies. In particular, the gas mass is obtained from the dust mass derived from the FIR SED of several hundreds AGN host galaxies at z<1, along with a control sample of normal galaxies selected in the same stellar mass and redshift ranges. We make use of a stacking procedure to increase the luminosity completeness of the studied samples (§ 5.2.2).

In § 5.1 we describe the selection of the AGN and inactive galaxy samples, in § 5.2 we derive the parameters we are interested in, especially the dust mass, and in 5.3 we discuss the results (included in Vito et al. 2014b).

5.1 Selection of the AGN and galaxy samples

5.1.1 Data set

We selected a sample of AGN at $z \leq 1$ (as well as normal galaxies for the control sample) in the COSMOS, GOODS-S and GOODS-N fields. The choice of these fields was driven by the huge multiwavelength coverage provided by a number of surveys,

which is crucial to derive reliable properties (redshift, stellar mass, star-formation rate and dust mass).

We used the Far-Infrared (FIR) data from the PACS Evolutionary Probe (PEP, Lutz et al. 2011) and the *Herschel* Multi-tiered Extra-galactic Survey (HerMES, Oliver et al. 2012) programs, which cover the three fields used in this work. The former was performed with the PACS camera (70, 100 and 160 μm; Poglitsch et al. 2010) while the latter with the SPIRE camera (250, 350 and 500 μm; Griffin et al. 2010), both on board of the *Herschel* Space Observatory (Pilbratt et al. 2010). *Herschel* catalogues are based on prior information on MIPS 24 μm positions and fluxes. The PEP catalogue was described by Lutz et al. (2011) and Berta et al. (2011), while the HerMES catalogue was presented by Roseboom et al. (2010; 2012). Since only the GOODS-S field was observed at 70 μm, following Santini et al. (2014), who reported that the exclusion of that band does not significantly affect the results, we will not use the 70 μm data, to use a consistent procedure among all the fields.

For the GOODS-S field we used the optical/near-IR photometric data from the GOODS-MUSIC catalogue (Grazian et al. 2006; Santini et al. 2009) and the X-ray counterpart information from the 4 Ms *Chandra* Deep Field South (CDF-S) main catalogue (Xue et al. 2011). In the GOODS-N field we collected the multiwavelength data from the PEP team catalogue (Berta et al. 2011) and the X-ray data from the 2 Ms *Chandra* Deep Field North catalogue (Alexander et al. 2003; Bauer et al. 2004). Finally, in the COSMOS field we used the Ilbert et al. (2009) and McCracken et al. (2010) multiwavelength catalogues and the *Chandra* (Civano et al. 2012) and XMM (Brusa et al. 2010) COSMOS optical identification catalogues, complemented by the photometric redshifts presented by Salvato et al. (2011). All the catalogues are supplemented with spectroscopic or photometric redshifts. Photometric redshift for inactive galaxies in COSMOS and GOODS-N lacking redshift information in the above-mentioned catalogues were computed by using the EAZY code (Brammer et al. 2008).

We applied two selection criteria to the parent sample in the three fields: 1) signal-to-noise ratio (SNR) ≥ 10 in the K-band; 2) $z \leq 1$. The K-band selection ensures to derive reliable estimates for the stellar mass (following Santini et al. 2014), while the redshift cut prevents the *Herschel* bands to be shifted to wavelengths that may be significantly contaminated by AGN heating and reprocessing (which may affect our estimate of the dust masses). Indeed, Rosario et al. (2012) have shown that out to z$\sim$1 the PACS colors are consistent with those typical of star forming galaxies, while some AGN contamination to the 100μm may occur at higher redshift.

5.1.2 AGN selection

We divided the selected objects into AGN and a galaxy samples. In COSMOS, we considered an X-ray detected object to be an AGN following either its optical classification (Brusa et al. 2010; Civano et al. 2012), its best-fitting SED template (Salvato et al. 2011) or if its absorption-corrected luminosity in the rest-frame $2-10$ keV band is $L_X \geq 10^{42}$ erg s^{-1}. This threshold is a compromise between the sample size and the contamination by purely powerful star-forming galaxies (Ranalli et al. 2003). The luminosities were collected from Lanzuisi et al. (2013) and Mainieri

et al. (2007; 2011), who performed a spectral analysis on a sample of bright sources in the *Chandra*-COSMOS (Elvis et al. 2009) and XMM-COSMOS (Cappelluti et al. 2009) catalogues, respectively. Intrinsic luminosities of all the XMM-COSMOS selected sources were also derived by the XMM-COSMOS team using the observed fluxes and Hardness Ratio HR (see Merloni et al. 2014). In GOODS-S we assumed the Xue et al. (2011) classification and luminosity for the X-ray sources. In GOODS-N we used the X-ray detected AGN sample of Bauer et al. (2004). The luminosity range of the X-ray selected AGN sample is $10^{41} \lesssim L_X < 10^{45}$ erg s^{-1}, and only ~ 5 per cent of them have $L_X > 10^{44}$ erg s^{-1}. Given the limiting fluxes of the X-ray samples, the most obscured AGN can still remain undetected. In order to at least partially recover them, we applied the IRAC power-law selection by Donley et al. (2012). The resulting AGN sample consists in 801 objects, of which 631 X-ray selected (486 in COSMOS and 145 in GOODS-S+N) and 255 IRAC selected (241 in COSMOS and 14 in GOODS-S+N). Eighty-five AGN (78 in COSMOS and 7 in GOODS-S+N) are in common between the two selections. All the sources which do not satisfy any of the AGN selection criterion in each field are automatically classified as inactive galaxies (~ 175000 objects).

5.2 Parameter derivation

In this section we describe how the physical parameters for our sample of AGN and galaxies were derived. Fig. 5.1 summarizes the basic steps of the procedure. In order to avoid selection effects due to redshift, stellar mass and luminosity affecting the dust mass (and hence the gas mass) measurement and distribution, we adopted a stacking technique to derive the average FIR fluxes in bins of stellar mass and redshift. This method ensures a high level of completeness, since only a small fraction of sources are individually detected by *Herschel* (e.g. $\sim 13\%$ of the AGN), especially in the longest wavelength bands.

5.2.1 Stellar mass

Stellar masses for AGN-hosting galaxies in the sample were derived using the same method as in Santini et al. (2012b), where several consistency tests were also performed. The observed optical photometry was fitted with a combined library of stellar synthetic templates (Bruzual & Charlot 2003) and pure AGN emission templates affected by levels of extinction (Silva et al. 2004) associated with different column densities N_H.

In order to break the degeneracies between the different fit parameters, the AGN were classified into Type I (unabsorbed) and Type II (absorbed) sources on the basis of optical classification (when available), best-fitting Spectral Energy Distribution (SED) template from Salvato et al. (2011) and intrinsic absorption derived from X-ray data analysis (Lanzuisi et al. 2013; Mainieri et al. 2007; 2011; Bauer et al. 2004). Sources for which all these priors are lacking, as well as X-ray undetected AGN (i.e. Spitzer selected AGN) were considered as Type II. Moreover, by averaging the two templates of moderate absorption, the four Silva et al. (2004) templates were reduced to three: unabsorbed ($N_H < 10^{22}$ cm^{-2}), Compton-thin absorbed ($10^{22} < N_H < 10^{24}$ cm^{-2}) and Compton-thick absorbed ($N_H > 10^{24}$ cm^{-2}). Only the most

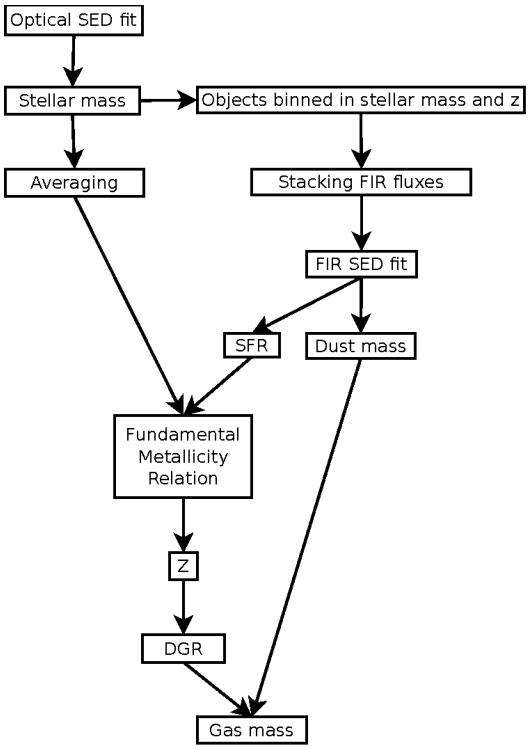

Figure 5.1: Schematic procedure followed to derive the gas mass.

suitable template, based on the above classification, was fitted to each AGN. In particular, the unabsorbed, absorbed Compton Thin and absorbed Compton-thick templates were used for 100, 694 and 7 sources, respectively.

The fit was performed through a χ^2 minimization up to $5\,\mu m$ rest-frame, since the Bruzual & Charlot (2003) templates do not include emission from dust reprocessing, assuming an exponentially declining star-formation history and a Salpeter IMF for the stellar component. Each flux was weighted by the inverse of the photometric error and the redshift of each object was fixed during the fit. The stellar mass was derived by the best-fitting stellar component alone.

5.2.2 Stacking procedure

AGN and galaxies are divided into 6 stellar mass bins (from $\log(M_*/M_\odot) = 9$ to 12) and 4 redshift bins. The number of AGN and galaxies in each bin is reported in Table 5.1. Hereafter, the stellar mass corresponding to each bin is assumed to be the mean stellar mass of the objects in the bin, with uncertainty equal to one standard deviation (typically ~ 0.14 dex for both AGN and galaxies). We adopt the stacking procedure implemented by Santini et al. 2014, where details on the stacking procedure are given (see also Shao et al. 2010, Rosario et al. 2012 and Santini et al. 2012b). However, in contrast with Santini et al., we do not bin in Star Formation Rate (SFR), since our goal is to investigate if differences in terms of gas content between AGNs and normal galaxies may actually be at the origin of their claimed SFR differences, hence we do not want to include any *a priori* selection or

Table 5.1: Number of AGN and galaxies (between brackets) stacked in each $z-$ M$_*$ bin (see § 5.2.2). The bins that fulfil the conditions in § 5.2.3 are in boldface.

z	log(M$_*$/M$_\odot$)					
	9-10	10-10.5	10.5-11	11-11.25	11.25-11.5	11.5-12
0.0-0.3	**10 (3120)**	< 10 (**735**)	**17 (496)**	< 10 (**125**)	< 10 (**37**)	< 10 (< 10)
0.3-0.6	**48 (13800)**	**32(3881)**	**65 (2795)**	**42 (801)**	**22 (414)**	**16 (117)**
0.6-0.8	29 (**10621**)	**24 (5065)**	**61 (3071)**	43 (**1018**)	**42 (504)**	16 (**199**)
0.8-1.0	17 (**6767**)	**19 (6209)**	**72 (4393)**	**70 (1372)**	**53 (754)**	**30 (295)**

binning on the SFR.

Here we briefly summarize the basic steps of the stacking procedure. For each *Herschel* band we excluded areas on the map where the integration time is lower than half the maximum, to avoid regions of high noise level. Then, for each $z - M_*$ bin containing at least 10 objects, we stack on the residual image (in which all the 3σ detected objects were subtracted) of each *Herschel* band at the position of the undetected sources in the bin, weighting by the inverse of the square of the error map. Fluxes on the stacked images were measured through a PSF fitting (for the PACS bands) or from the value of the central pixel (on the SPIRE images). Errors on the stacked fluxes were computed through a bootstrap procedure. Finally, the average flux S in each *Herschel* band is computed as:

$$S = \frac{S_{stack} \times N_{stack} + \sum_{i=1}^{N_{det}} S_i}{N_{stack} + N_{det}} \tag{5.1}$$

where S_{stack} is the stacked flux of the N_{stack} undetected objects in the bin and S_i is the flux of each of the N_{det} detected objects at the 3σ confidence level.

5.2.3 Dust Mass

FIR emission in star forming galaxies is mainly due to galactic cold dust, heated by young stars. The dust mass can be inferred by fitting dust emission models to the FIR fluxes varying the temperature distribution and normalization of the far-IR SED. We considered only $z-$ M$_*$ bins in which the stacked flux has a 3σ significance in at least three *Herschel* bands, at least one of which between the 350 and 500 μm bands. This requirement ensures a good sampling of the dust emission peak and, hence, reliable estimates of the dust mass.

We performed a χ^2 minimization to the analytical SED templates by Draine & Li (2007) (which span a broad range of dust temperature distributions) and, in each $z-$ M$_*$ bin, the best-fitting dust mass (M_d) was derived from the normalization of the template corresponding to the minimum χ^2. Further details on the method are given in Santini et al. (2014). Errors were estimated from the range of M_d covered by all the templates within $\Delta\chi^2 = 1$ from the value of the best fitting template. Regarding the AGN sample, since emission from AGN-heated dust could contribute to the FIR fluxes, especially in the PACS bands, we added the Silva et al. (2004) AGN SEDs during the Draine & Li (2007) templates fit, leaving their normalization free to vary. However, we shall mention that the results on the dust masses are essentially unchanged even by neglecting the AGN contribution in the SED fitting; this is not surprising, not only because AGN heating contribute mostly at mid-IR

wavelengths (which, at $z < 1$, are not relevant for the dust mass derivation), but also because most of the AGN in our sample do not have extreme luminosities (see § 5.1).

5.2.4 Gas mass

The total gas mass (M_{gas}, which incorporates both the molecular and atomic phases) can be inferred from the dust mass through a dust-to-gas ratio (DGR; e.g. Eales et al. 2010; Magdis et al. 2012): $M_{gas} = M_d /$ DGR. In order to derive the DGR, we assumed that a fixed fraction of metals is incorporated into dust grains (e.g. Draine et al. 2007 and references therein; Leroy et al. 2011; Smith et al. 2012; Corbelli et al. 2012; Sandstrom et al. 2012) and that the DGR scales linearly with the gas metallicity, traced by the oxygen abundance (Draine et al. 2007):

$$\mathrm{DGR} = 0.01 \cdot 10^{Z - Z_{\odot}}. \tag{5.2}$$

This assumption is observationally found to be a good approximation down to metallicities Z=12+log(O/H)∼8.0 (Rémy-Ruyer et al. 2013), which is certainly lower than the metallicity range spanned by our sample.

We derived the mean gas metallicity (Z) in each $z -$ M$_*$ bin from the Fundamental Metallicity Relation (FMR) by Mannucci et al. (2010; 2011), who found a narrow (scatter of ~ 0.05 dex) relation between M_*, SFR and Z in local and high-redshift galaxies.

The average SFR for each $z -$ M$_*$ bin was computed from the total 8–1000μm luminosity (LIR), derived by integration of the best-fitting dust emission template used in § 5.2.3:

$$\mathrm{SFR}[\mathrm{M}_{\odot}/\mathrm{yr}] = 1.8 \times 10^{-10}\, \mathrm{L_{IR}}[\mathrm{L}_{\odot}]. \tag{5.3}$$

Stellar mass and SFR were converted from a Salpeter to a Chabrier IMF (Davé 2008; Santini et al. 2012a), which was assumed by Mannucci et al. (2010; 2011). Errors on Z were estimated considering the FMR intrinsic scatter and the errors on M_*. Errors on SFR were ignored, since they are negligible with respect to the other uncertainties (typically ΔlogSFR/ logSFR $\simeq 0.04$). Finally, errors on M_d and Z were propagated to derive the uncertainties on M_{gas}. We shall mention that if the FMR is ignored and the metallicities are simply inferred from the mass-metallicity relation (Tremonti et al. 2004) and assuming a redshift evolution of this relation (Maiolino et al. 2008; Troncoso et al. 2014), the results remain unchanged. More generally, the bulk of the M$_{gas}$ variations are due to the variations of M$_{dust}$, while variations and uncertainties in the DGR play a secondary role (i.e. the main results do not change even if we only consider the dust mass instead of the gas mass, as we will discuss later on).

We also note that the effect of the presence of an AGN on the DGR, if any, would be to decrease its value with respect to inactive galaxies, since the hard nuclear radiation is expected to make the environment less favourable to dust survival. Nonetheless, we assumed the same DGR for AGN hosts as for inactive galaxies. As our results will show, this turns out to be a conservative approach.

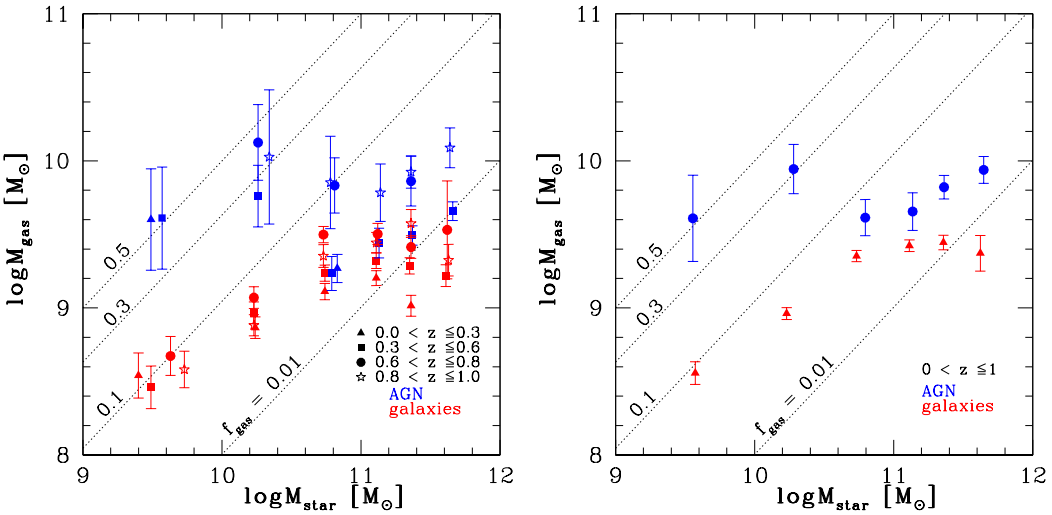

Figure 5.2: Gas mass as a function of the stellar mass for the AGN (blue symbols) and normal galaxy (red symbols) samples, in different redshift bins (identified by different symbols; left panel) and averaged over the redshift bins (right panel). Loci of constant gas fraction, for different values, are shown with dotted lines.

5.3 Results and discussion

Fig. 2.4 (left panel) shows the average gas mass as a function of the stellar mass of the AGN and galaxy samples for each $z-$ M$_*$ bin which fulfils the requirements described in § 5.2.3. The gas mass of galaxies increases with stellar mass, while the gas fraction, defined as $f_{gas} = M_{gas}/(M_{gas} + M_*)$, increases towards lower stellar masses, as already reported by Santini et al. (2014).

The most interesting result is that the AGN host gas masses are systematically higher than in the galaxy sample, in nearly all bins of stellar mass and redshift. Since this result does not show any evident dependence on redshift, in Fig. 5.2 (right panel) we plot the gas mass in the different stellar mass bins, averaged over the redshift bins. The same qualitative relations hold if one considers the directly-measured dust mass instead of the more physically relevant gas mass. Indeed, as also reported in § 5.2.4, gas mass derivation is strongly dominated by the dust mass, with variations on the DGR being second-order effects.

Fig. 5.3 shows the distribution of $M_{gas,AGN}/M_{gas,gal}$ for all redshift and stellar mass bins for which this comparison can be made. Clearly, in nearly all bins (more specifically, in 94% of them), $M_{gas,AGN}/M_{gas,gal} > 1$. In 44% of the bins the gas mass in AGN hosts is higher by more than a factor of three with respect to normal galaxies. The same result is shown in Fig. 5.4 in terms of gas fractions, $f_{gas,AGN}/f_{gas,gal}$.

To better quantify the significance of the result, Table 5.2 gives the mean $M_{gas,AGN}/M_{gas,gal}$ ratio, obtained by averaging the results in three different stellar mass ranges. The strongest difference is clearly at $\log(M_{star}/M_\odot) < 10.5$, where the gas mass of AGN hosts is on average an order of magnitude higher than in normal galaxies, with a significance of 6.5σ. However, the difference between

Table 5.2: Mean $\log(M_{gas,AGN}/M_{gas,gal})$ (computed in $z - M_*$ bins) for three different mass ranges.

$\log(\frac{M_*}{M_\odot})$	9-10.5	10.5-11.25	11.25-12
$\langle\log(\frac{M_{gas,AGN}}{M_{gas,gal}})\rangle$	1.04 ± 0.16	0.24 ± 0.08	0.44 ± 0.06

AGN and normal galaxies in terms of gas masses is significant at $\sim 3-7\sigma$ also at $\log(M_{star}/M_{odot}) > 10.5$. We checked that no significant change in the results are obtained considering different mass or redshift bins. The main conclusion holds, besides the loss of statistics, even repeating the analysis separately on X-ray detected and undetected AGN, as well as applying different cuts in L_X.

Since gas accretion onto SMBH is the process at the origin of nuclear activity, whatever the mechanisms that drives the gas into the central regions are, it is not surprising that AGN are preferentially hosted by gas rich galaxies (see also Silverman et al. 2009). Indeed, the conditions for gas accretion are statistically easier to be fulfilled in presence of a larger gas content: it is more likely that a gas cloud falls into the potential of the supermassive black hole if there are overall more gas clouds in the host galaxy. Beyond these simple statistical arguments, models have been proposed that ascribe AGN secular fuelling to disk instabilities (Bournaud et al. 2011), which are stronger in gas rich disks. Our results support this scenario.

As discussed above, the gas content is also the fundamental ingredient driving star formation in galaxies, through the SK relation. Several works have found a relation between strong nuclear activity and enhanced SFR with respect to inactive galaxies (e.g Lutz et al. 2008; 2010; Shao et al. 2010; Rosario et al. 2012). Rosario et al. (2013) and Santini et al. (2012b) found that the *Herschel* detection fraction for AGN is higher than for galaxies and concluded that AGN are more likely to be hosted by star forming galaxies.

Given the dependency of the SFR on the gas content and the result obtained by us, the enhanced star formation in AGN galaxies appears to be primarily the result of a larger (on average) gas content with respect to the bulk of the galaxy population (star forming and quiescent) at similar stellar masses, as already suggested by Rosario et al. (2012; 2013) and Santini et al. (2012b). However, the differences in terms of SFR are less clear with respect to the differences found by us in terms of gas content, probably because of the additional spread introduced by the SK relation, and the contribution of triggering mechanisms (e.g. galaxy interactions), which may affect the star formation efficiency per unit gas mass.

The main important result of our work is that AGN host galaxies are characterized by much larger amounts of gas, strongly suggesting that generally AGN activity in galaxies is simply fostered by a larger content of gas, without invoking specific triggering mechanisms, at least in the luminosity range probed by our work.

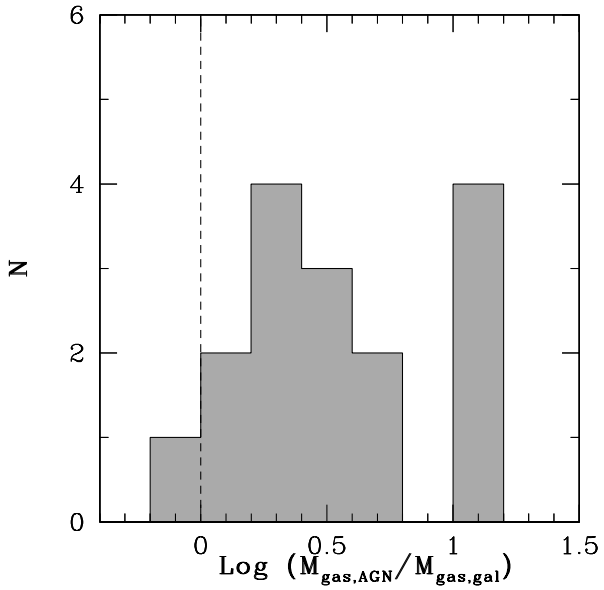

Figure 5.3: Distribution of the ratio between the gas mass in AGN hosts and in normal galaxies, inferred within the same $z-$ M_* bins.

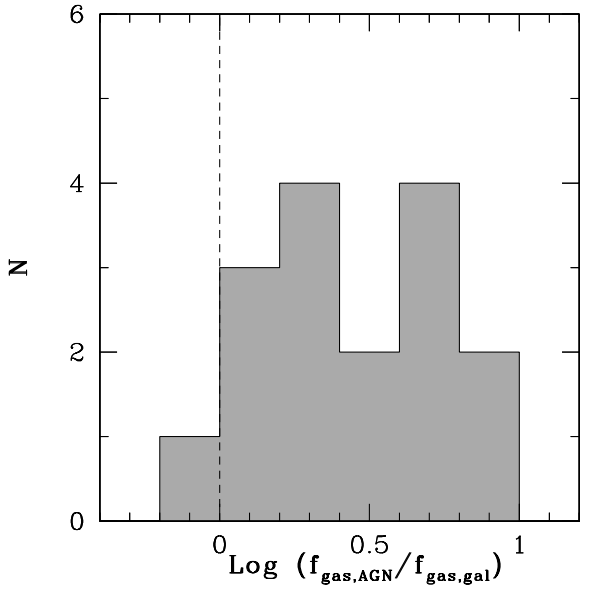

Figure 5.4: Distribution of the ratio between the gas fraction in AGN hosts and in normal galaxies, inferred within the same $z-$ M_* bins.

Chapter 6

Discussion and conclusions

In this thesis, we have investigated the evolution of the high-redshift ($z > 3$) AGN population. Observations in the X-rays are the best tools to achieve this goal, since X-rays are less affected by obscuration and galaxy dilution than optical/IR bands. We collected data from some of the major X-ray surveys, with different depths and areal coverage in order to sample as homogeneously as possible the $\log L_X - z$ plane. The 4 Ms CDFS allowed us to include faint Seyfert-like AGN ($L_X \sim 10^{43-44}\,\mathrm{erg\,s^{-1}}$), which constitute the bulk of the AGN population, while the rare luminous objects ($L_X > 10^{44}\,\mathrm{erg\,s^{-1}}$) are easier to be found in shallower and wider surveys like the *Chandra* and *XMM*-Newton observation in the COSMOS field and the SXDS. The selection of this sample took advantage of the most up-to-date spectroscopic and photometric redshift information in literature, in order to include a number as small as possible of low-redshift interlopers. The final sample (141 sources) is one of the largest selected in the X-rays at $z > 3$ and is characterised by a very high redshift completeness ($\sim 98\%$). We performed a spectral analysis on the *Chandra*-detected sources to derive the spectral parameters, while we relied on literature results for *XMM* sources.

We derived the hard X-ray luminosity functions (HXLF) at $z > 3$ of obscured and unobscured sources, by accounting for the proper spectral shape during the computation of the sky coverage. Then we added them to derive the total HXLF and fitted it with the most used evolutionary models in the literature applying a maximum likelihood procedure onto unbinned data. By restricting ourselves to the high-redshift domain, the results are not dominated by the behaviour at low redshift, where larger samples of AGN are easier to be collected and would dominate the fit. The best analytical representation of the AGN evolution at $z > 3$ is a pure density evolution (PDE) model: the AGN space density is found to decrease with increasing redshift as $(1+z)^{p}_{den}$, with $p_{den} \sim -6$, i.e. by a factor of ~ 10 from $z = 3$ to 5, confirming and expanding previous results (e.g. Hiroi et al. 2012), using a larger dataset and a different and more unbiased method. We did not find strong evidences in favour of a luminosity-dependent density evolution (LDDE). However, this could be simply hidden by the low statistics of low-luminosity $\log L_X < 44$ AGN, especially at $z > 4$, whose evolution would be the most affected by a LDDE-like evolution. If we assume our best-fitting values for a LDDE model, we are able to exclude an "upsizing" scenario, according to which the space density of low-luminosity AGN declines less steeply than that of luminous AGN with increasing

redshift, as suggested by Ueda et al. (2014). Although a luminosity and density evolution (LADE, Aird et al. 2010) model is formally consistent with our data, we found it to mimic the behaviour of the PDE model.

Hints in favour of a decline in the high-redshift AGN space density were also derived from the logN-logS which, thanks to the deep *Chandra* data, we were able to probe down to fluxes a factor of ~ 10 fainter than those achieved by most of the previous works. Compared to similar works performed in the 4 Ms CDFS, a different and somehow more conservative approach was assumed in our work. The number counts of our sample are consistent with the Gilli et al. (2007) XRB synthesis model, if the above-mentioned decline is included.

The results from the spectral analysis on sources selected in the 4 Ms CDFS were used to derive the intrinsic column-density distribution. The spectra were expected to suffer from an observational bias linked to the low photon-counting statistics which would lead to overestimate the intrinsic absorption level. Massive spectral simulations were performed to derive proper correction factors for that bias. After the correction, strongly-obscured ($\log N_H > 23$) sources represent the $\sim 57\%$ of the whole sample and $\sim 84\%$ of the most luminous ($\log L_X > 44$) objects. These values are larger than those reported at low redshift, in particular for luminous sources.

Having derived separately the HXLF for obscured and unobscured AGN, we could also give an estimate of the intrinsic obscured AGN fraction at $3 < z \lesssim 5$. This quantity was found to be constant with luminosity, around a value of ~ 0.5. Therefore, contrary to previous works performed at lower redshift, there is no hint of an anti-correlation between the obscured AGN fraction and the luminosity. Comparing our data with results derived in the Local Universe by Burlon et al. (2011), we find a positive evolution of the obscured AGN fraction with redshift, especially for luminous ($\log L_X > 44$) AGN. The rather narrow redshift range probed by this work prevented us from drawing statistically significant results on the evolution of the obscured AGN fraction from $z = 3$ to 5. However, we note that Hasinger (2008) suggested that this quantity saturates at $z > 2$.

These findings fit well in a wider scenario in which the evolution of obscuration is related to the triggering mechanism of the nuclear activity and is thought to be different for low and high-luminosity AGN. Indeed, the activity of luminous ($L_{2-10\ \mathrm{keV}} \geq 10^{44}$ erg s^{-1}) AGN, which occurs in short bursts (with timescales of the order of ~ 0.01 Gyr; e.g. Alexander et al. 2005), is likely to be triggered by major gas-rich mergers. In this case, the gas accretion is chaotic (Hopkins et al. 2008) and may produce a large covering factor and column densities. Since the merging rate and, most importantly, the gas fraction increase with redshift, the fraction of obscured AGN may then increase as well. At lower luminosities, the accretion onto the SMBH is probably driven by secular processes (~ 1 Gyr; e.g. Daddi et al. 2007; Elbaz et al. 2011) and is expected to be smoother and symmetrical. Therefore, the fraction of obscured AGN would be constant with redshift, given that obscuration would be a function of the system geometry only. Since a threshold of $\log N_H = 23$ was used to divide obscured and unobscured AGN, a even larger obscured AGN fraction that the value reported in this thesis (~ 0.5) would be derived if the widely used threshold $\log N_H = 22$ would be used. Therefore, this result suggests that at high-redshift the majority of AGN are obscured.

To better assess the evolution of the AGN population and the obscured AGN

fraction at high-redshift, larger samples of AGN, especially of low luminosity (down to $\log L_X \sim 42$), are needed. In the near future they will be provided by new X-ray surveys, like the *Chandra*-COSMOS Legacy Survey (PI: F. Civano) and the additional 3 Ms *Chandra* observations in the CDFS (PI: N. Brandt), which are expected to detect hundreds and tens of $z > 3$ AGN, respectively. The high-luminosity regime could be better investigated by including data from shallow X-ray surveys like *XMM*/XXL and the future eRosita All Sky Survey, which are expected to detect hundreds of high-redshift AGN. I plan to improve and expand this work in the future by including the new sources which will be provided by such surveys. The improved statistics will allow me to better estimate the parameters of the HXLF at both the faint and bright end. This will be crucial to constrain the evolution of the $z > 3$ AGN population.

I will also perform a stacking analysis of the X-ray emission of high-redshift ($z \sim 4 - 8$) AGN candidates and galaxies selected as drop-out and/or Lyman-α Emitter galaxies, e.g. in the CANDELS observations. Approximately 3000 drop-out galaxies at $4 < z < 9$ selected in the ultra-deep *Hubble* fields in the CDFS are included in the most recent compilations (e.g. Bouwens et al. 2014). I will also use of photometric redshifts provided by several papers in literature to refine such number. During the stacking process, massive spectral simulations will be necessary to properly interpret the results derived from the average signal. The analysis will allow me to study the average X-ray spectral parameters such as intrinsic column density and luminosity of accreting black holes as a function of redshift and to possibly put constraints on black hole seed models. In fact, I will sample the medium-to-low luminosity regime of AGN activity, characteristic of the bulk of galaxies, in the early stages of the Universe. Therefore, this investigation will provide information on the early BH growth complementary to that from studies of individual, very luminous quasars (e.g. Moretti et al. 2014).

In the current evolutionary scheme for AGN, the gas obscuring the nuclear regions may be linked to that responsible for star formation. In this framework, the study of the gas content of high-redshift AGN would return interesting constraints on the BH-galaxy co-evolution (e.g. Gilli et al. 2014). Unfortunately, a systematic study at high redshift is not feasible with current data. Therefore, in this thesis, we studied the total (molecular and atomic) gas content of AGN-hosting galaxies and compared it with that of inactive galaxies, limiting ourselves to $z < 1$. For the first time, we applied to an AGN sample a method to derive the gas mass previously used for inactive galaxies only. The method is based on the derivation of the dust mass of a galaxy from the fit of its FIR SED. To this goal, we selected in the K-band a sample of AGN and a control sample of inactive (starforming and quiescent) galaxies in the COSMOS, GOODS-N and GOODS-S fields, using the huge multiwavelength data provided by several surveys. We limited to $z < 1$ because at higher redshift the AGN emission would possibly strongly contaminate the FIR bands, affecting the dust mass estimate. The objects were divided in bins of stellar mass and redshift and a stacking procedure was applied on all the sources in each bin in order to derive the average FIR fluxes for AGN and galaxies. These fluxes were then fitted to derive the average dust mass, which was finally converted into a gas mass. The stacking procedure allowed us to include also sources not individually detected in the FIR bands and therefore to increase the completeness of the sample. By comparing the

average gas content of AGN and normal galaxies in the same bins of stellar mass and redshift, we found that the gas mass in AGN-hosting galaxies is systematically larger than in the bulk of inactive galaxies, i.e. AGN preferentially live in gas-rich galaxies. This result can be understood in simple statistical terms: it is easier for a gas cloud to fall into the nuclear region and trigger the nuclear activity if there are overall more gas clouds in the galaxy. This means that the specific triggering mechanism (merger, bars, etc) likely plays a secondary role in activating an AGN, with the presence of a large amount of gas being the fundamental requirement for that process. However, disk instabilities, which are often thought to be important for driving the gas toward the central region of a galaxy and therefore facilitating the triggering of AGN of low and moderate luminosity, grow stronger in gas-rich disks. Our result also naturally explains why AGN are preferentially hosted by star-forming galaxies, since a large amount of gas is essential to produce episodes of star formation.

The natural follow-up of this work would be to include sub-mm/mm data (e.g. from ALMA observations in COSMOS, see Scoville et al. 2014) during the FIR SED fitting. Sub-mm/mm data, which are not severely contaminated by AGN emission even at $z > 1$, would allow me on the one hand to derive more reliable dust, hence gas, masses, on the other hand to extend the investigation up to higher ($z \sim 1 - 3$) redshifts, where both the black hole accretion rate density and the star formation density peak. Assessing the gas content in AGN-hosting galaxies in that redshift range is crucial to constrain the AGN-star formation link.

Appendix A

Atlas of X-ray spectra

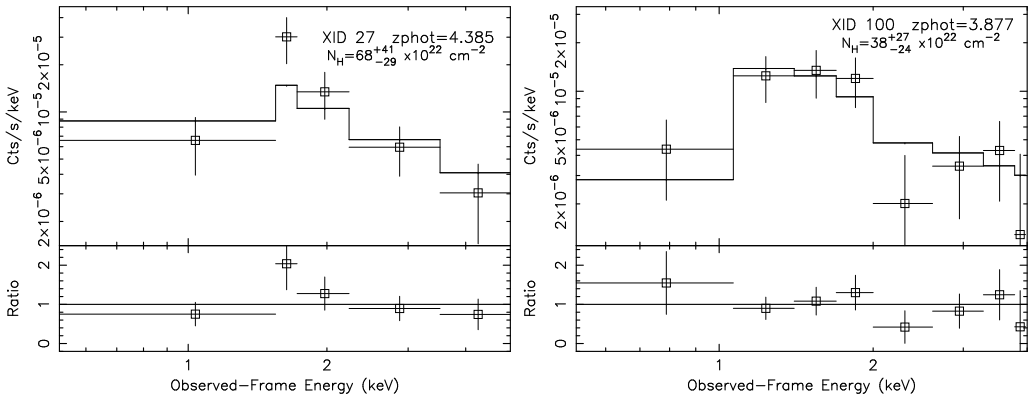

Figure A.1: X-ray spectra of the 34 sources in the high-redshift sample in the CDFS, fitted with an absorbed power-law model with $\Gamma = 1.8$ (see § 2.3.2). The ratio of data to folded model is shown and the spectrum is binned (at $\geq 3\sigma$ detection significance) for display purposes only. Identification numbers are from Xue et al. (2011). The spectroscopic or photometric redshift is reported for each source, as well as the best-fitting column density, accordingly to Tab. 2.1. (Continue)

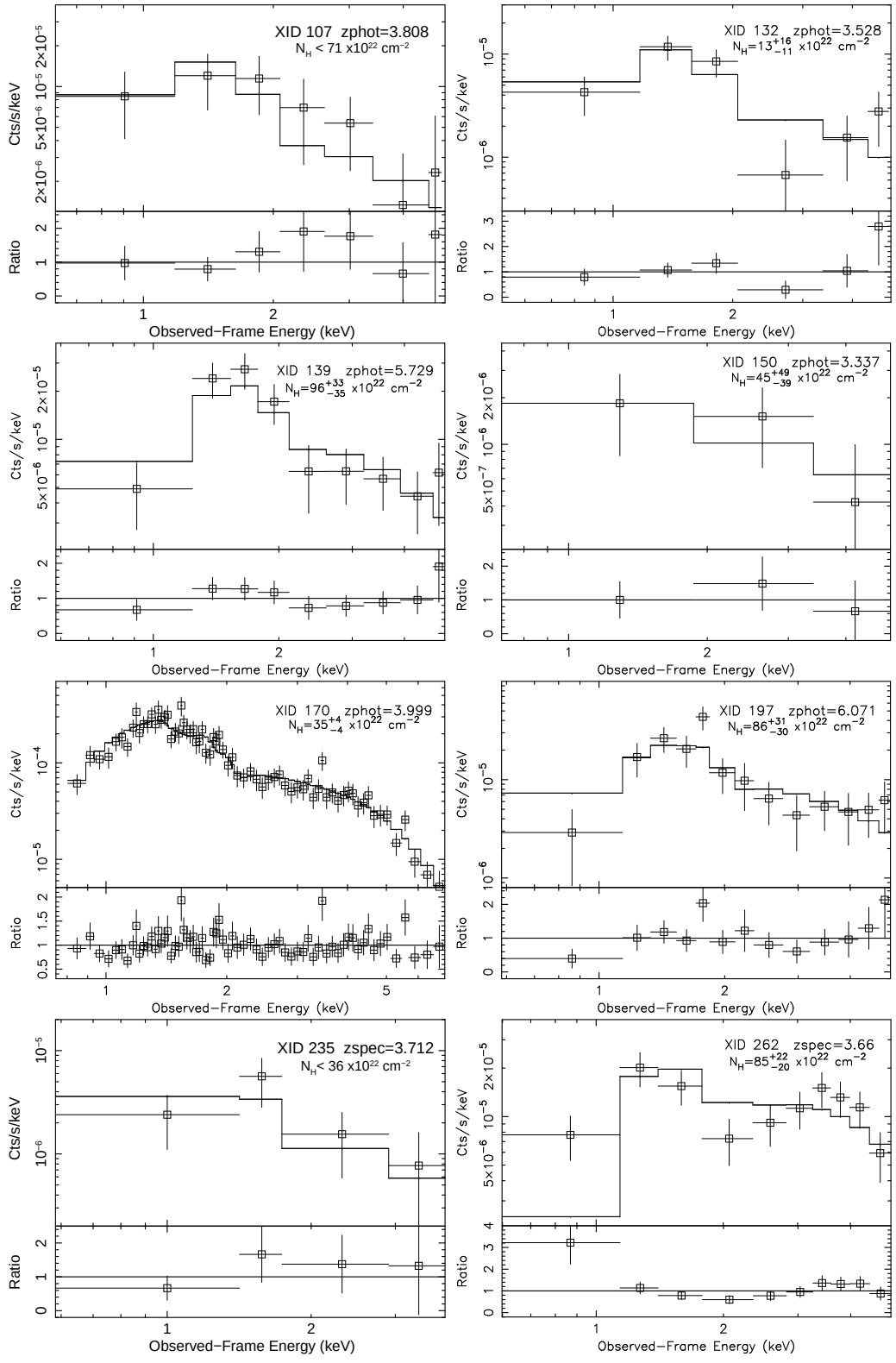

Figure A.1: (Continue)

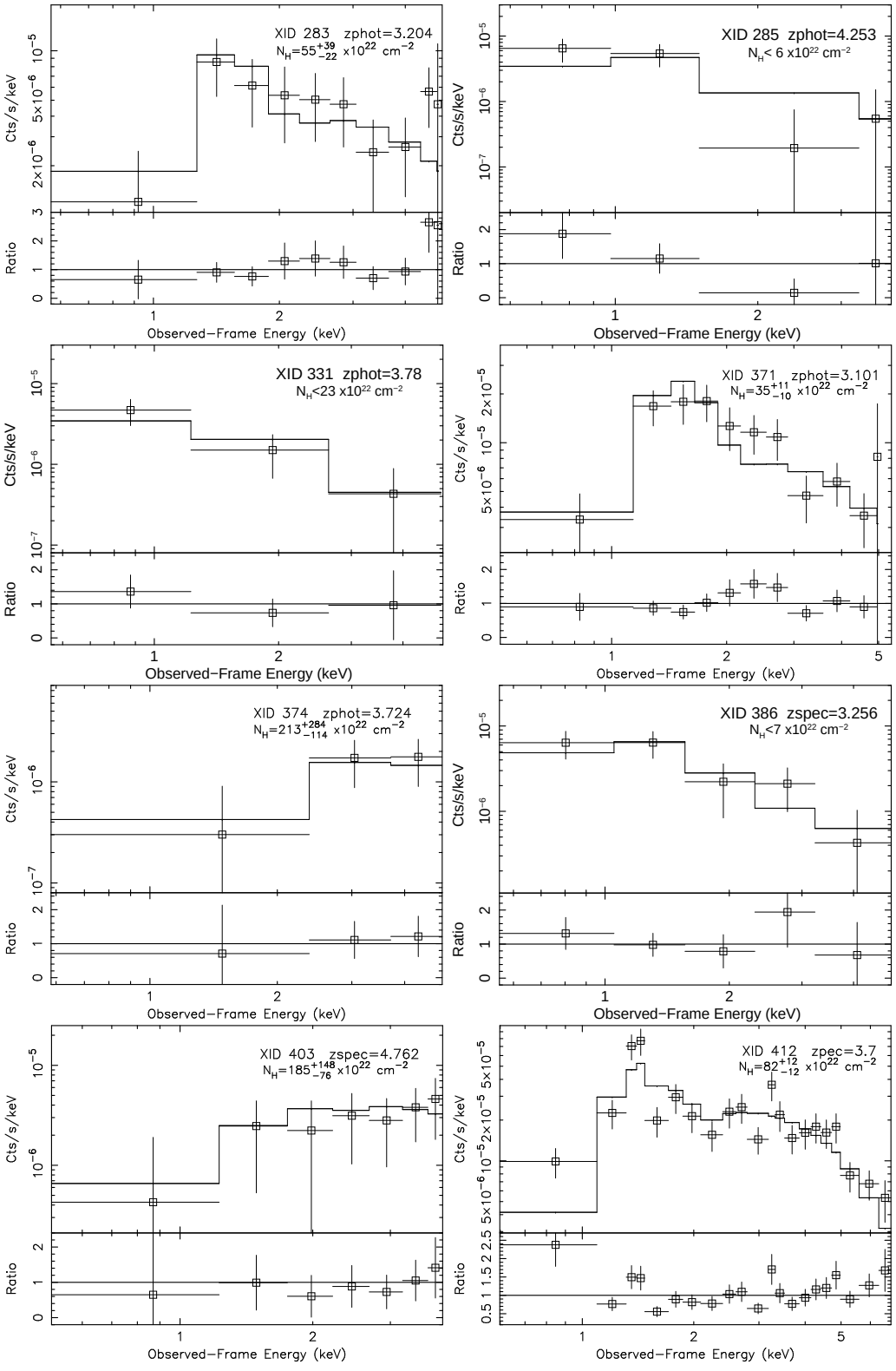

Figure A.1: (Continue)

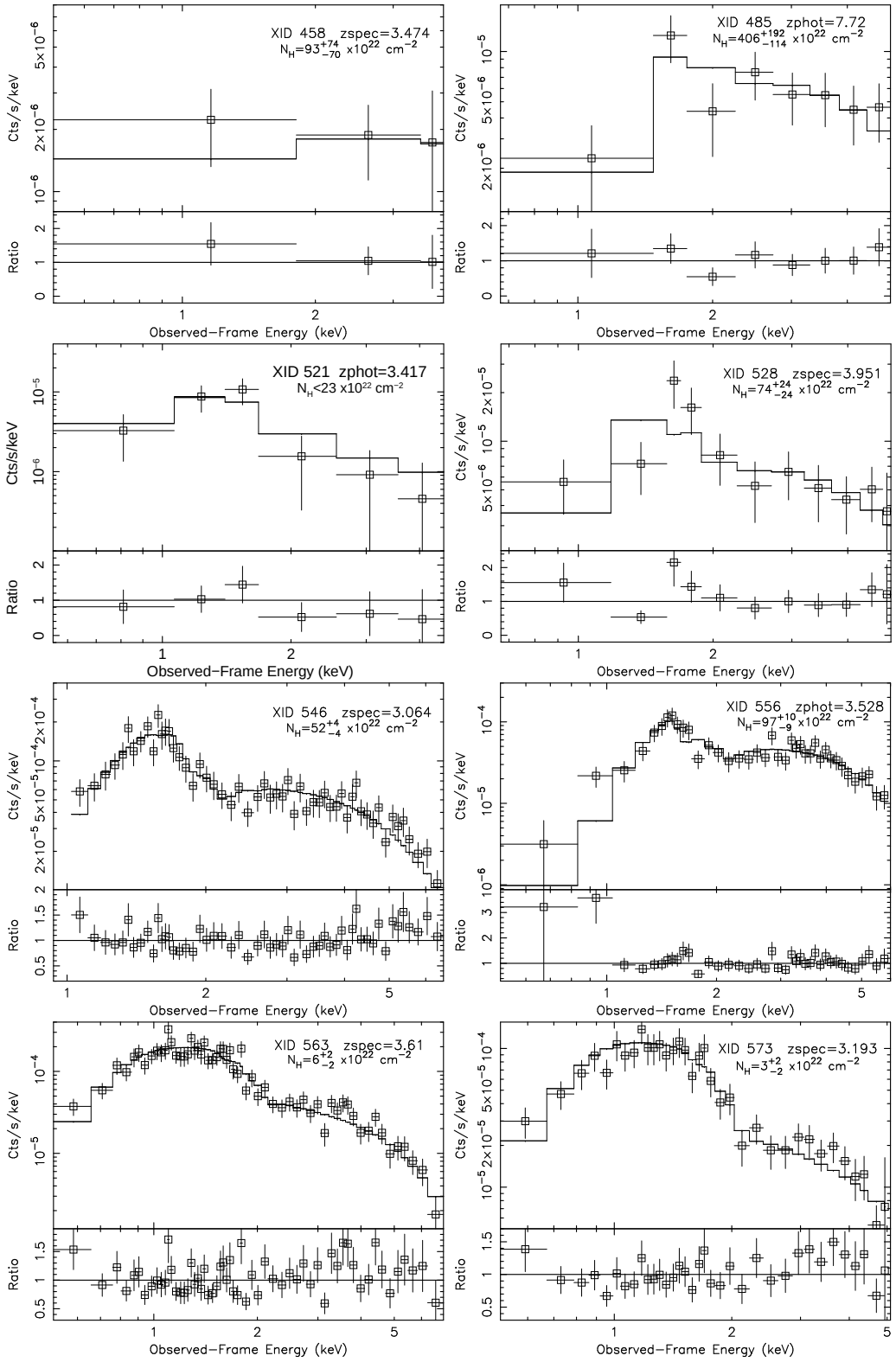

Figure A.1: (Continue)

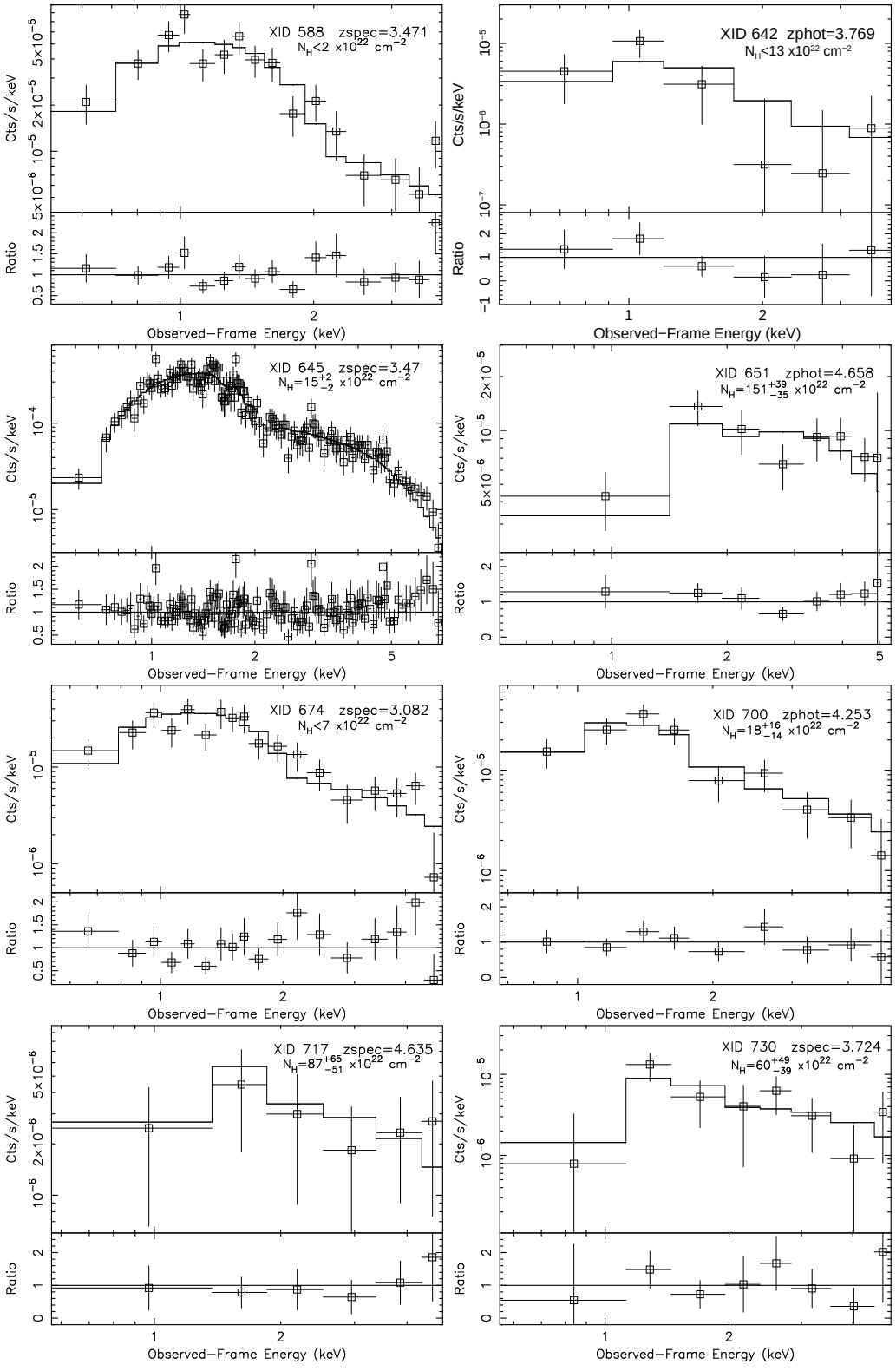

Figure A.1: (Continue)

Appendix B

Deriving the correction factors for the observed N_H distribution from simulations of X-ray spectra

The correction factors discussed in § 2.3.3 are derived from spectral-fitting simulations using XSPEC, assuming five different degrees of obscuration (see § 2.3.3 and Tab. 2.4). The five cases are simulated by runs of 1000 simulations each. The spectra are simulated at $z = 4$ with an input normalization tuned to obtain a distribution of net counts peaked at ≈ 100 (with a resulting dispersion of $\sigma \approx 11\%$), using the response files of a real source at an off-axis angle $\theta = 4$ arcmin. The results do not vary significantly if different response files are used, but they are sensitive to the number of net counts and redshift. To be consistent with the real cases, for these parameters we assumed values close to the average ones of the sample.

The simulation procedure can be summarized in the following steps:

1. Definition of an input model: we considered a power-law with photon index fixed to $\Gamma = 1.8$ as input model, absorbed by five different column densities, whose values vary between the five considered cases (Tab. 2.4, third column). Galactic absorption ($N_H = 7 \times 10^{19}$ cm^{-2}) is also included. The desired number of counts is obtained by properly adjusting the normalization of the power-law.

2. Simulation of the spectrum: the *fakeit* command in XSPEC can simulate a spectrum with background, given a starting model. It requires as input the response and ancillary files, a real background file (with the POISSERR keyword set to TRUE) and the desired exposure time. We set it to 3.5×10^6 s to be consistent with the average real case. The output consists of a fake source and a fake background spectrum file. The same BACKSCAL keyword, read from the header of the input background file, is assigned by the *fakeit* procedure to both simulated spectra, which are therefore assumed to be extracted from regions with an equal area. This approach is not correct and causes the source and background counts to be wrongly scaled (i.e. a

strong underestimate of the source net-count rate). To address this problem, before simulating the spectra we multiplied the EXPOSURE keyword of the real background file by a factor equal to the ratio between the BACKSCAL parameters of the real background and associated source file. We checked that the errors on the net-count rate were correctly computed (i.e. propagating the errors of the total and background-only count-rates).

3. Spectral fitting: after having grouped the simulated spectra, using the FTOOL *grppha*, with at least 1 count per bin, and assigned to them the response files and the fake background spectra, the spectra are repeatedly fit in the energy range E = 0.5 − 5 keV with an absorbed power-law model with N_H free to vary and photon index fixed to $\Gamma = 1.8$, accounting also for the Galactic absorption (fixed to $N_H = 7 \times 10^{19}$ cm^{-2}). Cash statistics were employed in the fitting. The fitting step (which use a local minimization algorithm) is alternated to the computation of the errors of the various free parameters (which is a more global operation) for a few times, in order to avoid the fit to be stuck in a local minimum.

For each of the five considered cases of input N_H, we obtained as output 1000 best-fitting values of N_H. Following the procedure adopted to fit the real spectra, we define N_H to be constrained if the lower limit (at the 90 per cent c.l.) on the best-fitting value is larger than zero, otherwise we consider its upper limit (e.g. if the fit returns a best-fitting value $N_H = 10^{+12}_{-10} \times 10^{22}$ cm^{-2}, we consider $N_H < 22 \times 10^{22}$ cm^{-2}).

The probability p_{ij} that a simulated source in the j-th N_H bin is observed in the i-th N_H bin, are derived by counting the number of times a constrained best-fitting N_H falls in bin i and normalized this number to 1000. In order to conserve the number of sources when the correction factors are applied on the real distribution, the p_{ij} factors must be rescaled to the total probability to constrain N_H (i.e. we excluded all cases in which only upper limits could be found):

$$P_{ij} = \frac{p_{ij}}{\sum_{i=A}^{E} p_{ij}} \tag{B.1}$$

As discussed in § 2.3.3, bins A, B and C are merged into a single bin (ABC), since they are indistinguishable at $z > 3$. Therefore, the probability factors, $P_{i\ (\mathrm{ABC})}$, are:

$$P_{\ (\mathrm{ABC})(\mathrm{ABC})} = \sum_{i} \frac{\sum_j p_{ij}}{3} \tag{B.2}$$

for i, j = A, B, C; $P_{i\ (\mathrm{ABC})} = \frac{p_{i\ \mathrm{A}} + p_{i\ \mathrm{B}} + p_{i\ \mathrm{C}}}{3}$ and $P_{\ (\mathrm{ABC})j} = \frac{p_{\ \mathrm{A}j} + p_{\ \mathrm{B}j} + p_{\ \mathrm{C}j}}{3}$ for i, j = D and E.

The same procedure is then repeated using $\Gamma = 1.6$ instead of $\Gamma = 1.8$. The P_{ij} factors, for i and j = (ABC), D and E, are those used in § 2.3.3.

Bibliography

Adelman-McCarthy J. K. et al., 2006, ApJS, 162, 38

Aird J. et al., 2010, MNRAS, 401, 2531

Akylas A., Georgakakis A., Georgantopoulos I., Brightman M., Nandra K., 2012, A&A, 546, A98

Akylas A., Georgantopoulos I., Georgakakis A., Kitsionas S., Hatziminaoglou E., 2006, A&A, 459, 693

Alexander D. M. et al., 2003, AJ, 126, 539

Alexander D. M., Bauer F. E., Chapman S. C., Smail I., Blain A. W., Brandt W. N., Ivison R. J., 2005, ApJ, 632, 736

Alvarez M. A., Wise J. H., Abel T., 2009, ApJ, 701, L133

Antonucci R., 1993, ARAA, 31, 473

Arnaud K. A., 1996, in Astronomical Society of the Pacific Conference Series, Vol. 101, Astronomical Data Analysis Software and Systems V, Jacoby G. H., Barnes J., eds., p. 17

Assef R. J. et al., 2013, ApJ, 772, 26

Avni Y., 1976, ApJ, 210, 642

Avni Y., Bahcall J. N., 1980, ApJ, 235, 694

Balestra I. et al., 2010, A&A, 512, A12

Barger A. J., Cowie L. L., 2005, ApJ, 635, 115

Bauer F. E., Alexander D. M., Brandt W. N., Schneider D. P., Treister E., Hornschemeier A. E., Garmire G. P., 2004, AJ, 128, 2048

Beckwith S. V. W. et al., 2006, AJ, 132, 1729

Begelman M. C., 2010, MNRAS, 402, 673

Begelman M. C., Volonteri M., Rees M. J., 2006, MNRAS, 370, 289

Bennert V. N., Auger M. W., Treu T., Woo J.-H., Malkan M. A., 2011, ApJ, 742, 107

Berta S. et al., 2011, A&A, 532, A49

Bianchi S., Guainazzi M., Matt G., Fonseca Bonilla N., 2007, A&A, 467, L19

Bianchi S., Maiolino R., Risaliti G., 2012, Advances in Astronomy, 2012

Blackburn J. K., 1995, in Astronomical Society of the Pacific Conference Series, Vol. 77, Astronomical Data Analysis Software and Systems IV, Shaw R. A., Payne H. E., Hayes J. J. E., eds., p. 367

Bolatto A. D., Wolfire M., Leroy A. K., 2013, ARA&A, 51, 207

Bondi M., Ciliegi P., Schinnerer E., Smolčić V., Jahnke K., Carilli C., Zamorani G., 2008, ApJ, 681, 1129

Bongiorno A. et al., 2012, MNRAS, 427, 3103

Bongiorno A. et al., 2010, A&A, 510, A56

Bournaud F., Dekel A., Teyssier R., Cacciato M., Daddi E., Juneau S., Shankar F., 2011, ApJ, 741, L33

Bouwens R. J. et al., 2014, ApJ, 793, 115

Boyle B. J., Terlevich R. J., 1998, MNRAS, 293, L49

Brammer G. B., van Dokkum P. G., Coppi P., 2008, ApJ, 686, 1503

Brammer G. B. et al., 2012, ApJS, 200, 13

Brandt W. N., Hasinger G., 2005, ARA&A, 43, 827

Bromm V., Loeb A., 2003, ApJ, 596, 34

Brusa M. et al., 2010, ApJ, 716, 348

Brusa M. et al., 2009, ApJ, 693, 8

Bruzual G., Charlot S., 2003, MNRAS, 344, 1000

Burlon D., Ajello M., Greiner J., Comastri A., Merloni A., Gehrels N., 2011, ApJ, 728, 58

Caldwell J. A. R. et al., 2008, ApJS, 174, 136

Cano-Díaz M., Maiolino R., Marconi A., Netzer H., Shemmer O., Cresci G., 2012, A&A, 537, L8

Capak P. et al., 2007, ApJS, 172, 99

Cappelluti N. et al., 2009, A&A, 497, 635

Cardamone C. N. et al., 2010, ApJS, 189, 270

Carilli C. L., Walter F., 2013, ARA&A, 51, 105

Cash W., 1979, ApJ, 228, 939

Chiaberge M., Celotti A., Capetti A., Ghisellini G., 2000, A&A, 358, 104

Cicone C. et al., 2014, A&A, 562, A21

Civano F. et al., 2011, ApJ, 741, 91

Civano F. et al., 2012, ApJS, 201, 30

Comastri A. et al., 2011, A&A, 526, L9

Comastri A., Setti G., Zamorani G., Hasinger G., 1995, A&A, 296, 1

Corbelli E. et al., 2012, A&A, 542, A32

Cowie L. L., Songaila A., Hu E. M., Cohen J. G., 1996, AJ, 112, 839

Daddi E. et al., 2007, ApJ, 670, 173

Dahlen T. et al., 2010, ApJ, 724, 425

Damen M. et al., 2011, ApJ, 727, 1

Davé R., 2008, MNRAS, 385, 147

Della Ceca R. et al., 2008, Mem. Soc. Astron. Italiana, 79, 65

Delvecchio I. et al., 2014, MNRAS, 439, 2736

Devecchi B., Volonteri M., 2009, ApJ, 694, 302

Di Matteo T., Springel V., Hernquist L., 2005, Nature, 433, 604

Donley J. L. et al., 2012, ApJ, 748, 142

Draine B. T. et al., 2007, ApJ, 663, 866

Draine B. T., Li A., 2007, ApJ, 657, 810

Dwelly T., Page M. J., 2006, MNRAS, 372, 1755

Eales S. A. et al., 2010, A&A, 518, L62

Ebrero J. et al., 2009, A&A, 493, 55

Eisenstein D. J., Loeb A., 1995, ApJ, 443, 11

Elbaz D. et al., 2011, A&A, 533, A119

Elvis M. et al., 2009, ApJS, 184, 158

Fabian A. C., Iwasawa K., Reynolds C. S., Young A. J., 2000, PASP, 112, 1145

Fabian A. C., Miniutti G., 2005, ArXiv Astrophysics e-prints

Fan X. et al., 2006a, AJ, 132, 117

Fan X. et al., 2006b, AJ, 131, 1203

Fasano G., Franceschini A., 1987, MNRAS, 225, 155

Ferrarese L., Merritt D., 2000, ApJ, 539, L9

Feruglio C., Maiolino R., Piconcelli E., Menci N., Aussel H., Lamastra A., Fiore F., 2010, A&A, 518, L155

Fiore F. et al., 2012, A&A, 537, A16

Fischer J. et al., 2010, A&A, 518, L41

Fontanot F., De Lucia G., Monaco P., Somerville R. S., Santini P., 2009, MNRAS, 397, 1776

Furusawa H. et al., 2008, ApJS, 176, 1

Gehrels N., 1986, ApJ, 303, 336

Giacconi R., Zamorani G., 1987, ApJ, 313, 20

Giacconi R. et al., 2002, ApJS, 139, 369

Giavalisco M. et al., 2004, ApJ, 600, L93

Gilli R., Comastri A., Hasinger G., 2007, A&A, 463, 79

Gilli R., Comastri A., Vignali C., Ranalli P., Iwasawa K., 2010, X-ray Astronomy 2009; Present Status, Multi-Wavelength Approach and Future Perspectives, 1248, 359

Gilli R. et al., 2014, A&A, 562, A67

Gilli R., Salvati M., Hasinger G., 2001, A&A, 366, 407

Gilli R. et al., 2011, ApJ, 730, L28

Grazian A. et al., 2006, A&A, 449, 951

Green P. J. et al., 2009, ApJ, 690, 644

Griffin M. J. et al., 2010, A&A, 518, L3

Grogin N. A. et al., 2011, ApJS, 197, 35

Guainazzi M., Bianchi S., Dovčiak M., 2006, Astronomische Nachrichten, 327, 1032

Guo Y. et al., 2013, ApJS, 207, 24

Haardt F., Maraschi L., 1993, ApJ, 413, 507

Haehnelt M. G., Rees M. J., 1993, MNRAS, 263, 168

Häring N., Rix H.-W., 2004, ApJ, 604, L89

Harrison C. M., Alexander D. M., Mullaney J. R., Swinbank A. M., 2014, MNRAS, 441, 3306

Hasinger G., 2008, A&A, 490, 905

Hasinger G. et al., 2007, ApJS, 172, 29

Hasinger G., Miyaji T., Schmidt M., 2005, A&A, 441, 417

Hickox R. C. et al., 2009, ApJ, 696, 891

Hiroi K., Ueda Y., Akiyama M., Watson M. G., 2012, ApJ, 758, 49

Hopkins A. M., Beacom J. F., 2006, ApJ, 651, 142

Hopkins P. F., Hernquist L., Cox T. J., Kereš D., 2008, ApJS, 175, 356

Ilbert O. et al., 2009, ApJ, 690, 1236

Iwasawa K. et al., 2012, A&A, 546, A84

Iwasawa K., Taniguchi Y., 1993, ApJ, 413, L15

Jahnke K., Macciò A. V., 2011, ApJ, 734, 92

Jiang P., Wang J. X., Wang T. G., 2006, ApJ, 644, 725

Joye W. A., Mandel E., 2003, in Astronomical Society of the Pacific Conference Series, Vol. 295, Astronomical Data Analysis Software and Systems XII, Payne H. E., Jedrzejewski R. I., Hook R. N., eds., p. 489

Kalberla P. M. W., Burton W. B., Hartmann D., Arnal E. M., Bajaja E., Morras R., Pöppel W. G. L., 2005, A&A, 440, 775

Kalfountzou E., Civano F., Elvis M., Trichas M., Green P., 2014, ArXiv e-prints

Kartaltepe J. S. et al., 2010, ApJ, 709, 572

Kartaltepe J. S. et al., 2007, ApJS, 172, 320

Kauffmann G. et al., 2003, MNRAS, 346, 1055

Kennicutt J. R. C., 1998, ApJ, 498, 541

Kim M. et al., 2007, ApJS, 169, 401

Koekemoer A. M. et al., 2011, ApJS, 197, 36

Koratkar A., Blaes O., 1999, PASP, 111, 1

Kormendy J., Ho L. C., 2013, ARA&A, 51, 511

Koushiappas S. M., Bullock J. S., Dekel A., 2004, MNRAS, 354, 292

La Franca F. et al., 2005, ApJ, 635, 864

La Franca F. et al., 2002, ApJ, 570, 100

Lagos C. D. P., Baugh C. M., Lacey C. G., Benson A. J., Kim H.-S., Power C., 2011, MNRAS, 418, 1649

Lamastra A., Perola G. C., Matt G., 2006, A&A, 449, 551

Lanzuisi G. et al., 2013, MNRAS, 431, 978

Lawrence A., 1991, MNRAS, 252, 586

Lawrence A., Elvis M., 2010, ApJ, 714, 561

Le Fèvre O. et al., 2004, A&A, 428, 1043

Le Floc'h E. et al., 2009, ApJ, 703, 222

Lehmer B. D. et al., 2012, ApJ, 752, 46

Leroy A. K. et al., 2011, ApJ, 737, 12

Lilly S. J. et al., 2009, ApJS, 184, 218

Lilly S. J. et al., 2007, ApJS, 172, 70

Lodato G., Natarajan P., 2006, MNRAS, 371, 1813

Loeb A., Rasio F. A., 1994, ApJ, 432, 52

Luo B. et al., 2008, ApJS, 179, 19

Luo B. et al., 2010, ApJS, 187, 560

Lusso E. et al., 2013, ApJ, 777, 86

Lutz D. et al., 2010, ApJ, 712, 1287

Lutz D. et al., 2011, A&A, 532, A90

Lutz D. et al., 2008, ApJ, 684, 853

Madau P., Dickinson M., 2014, ArXiv e-prints

Madau P., Rees M. J., 2001, ApJ, 551, L27

Magdis G. E. et al., 2012, ApJ, 760, 6

Magdis G. E. et al., 2011, ApJ, 740, L15

Magorrian J. et al., 1998, AJ, 115, 2285

Mainieri V. et al., 2011, A&A, 535, A80

Mainieri V. et al., 2007, ApJS, 172, 368

Maiolino R. et al., 2012, MNRAS, 425, L66

Maiolino R. et al., 2008, A&A, 488, 463

Maiolino R., Shemmer O., Imanishi M., Netzer H., Oliva E., Lutz D., Sturm E., 2007, A&A, 468, 979

Mannucci F., Cresci G., Maiolino R., Marconi A., Gnerucci A., 2010, MNRAS, 408, 2115

Mannucci F., Salvaterra R., Campisi M. A., 2011, MNRAS, 414, 1263

Marconi A., Hunt L. K., 2003, ApJ, 589, L21

Marconi A., Risaliti G., Gilli R., Hunt L. K., Maiolino R., Salvati M., 2004, MNRAS, 351, 169

Marshall H. L., Tananbaum H., Avni Y., Zamorani G., 1983, ApJ, 269, 35

Matute I., La Franca F., Pozzi F., Gruppioni C., Lari C., Zamorani G., 2006, A&A, 451, 443

McConnell N. J., Ma C.-P., 2013, ApJ, 764, 184

McCracken H. J. et al., 2010, ApJ, 708, 202

Menci N., Fiore F., Puccetti S., Cavaliere A., 2008, ApJ, 686, 219

Merloni A. et al., 2014, MNRAS, 437, 3550

Mignoli M. et al., 2005, A&A, 437, 883

Mignoli M. et al., 2013, A&A, 556, A29

Miller N. A., Fomalont E. B., Kellermann K. I., Mainieri V., Norman C., Padovani P., Rosati P., Tozzi P., 2008, ApJS, 179, 114

Miniutti G., Piconcelli E., Bianchi S., Vignali C., Bozzo E., 2010, MNRAS, 401, 1315

Miyaji T., Hasinger G., Schmidt M., 2000, A&A, 353, 25

Mor R., Trakhtenbrot B., 2011, ApJ, 737, L36

Moretti A. et al., 2014, A&A, 563, A46

Mortlock D. J. et al., 2011, Nature, 474, 616

Nenkova M., Sirocky M. M., Ivezić v., Elitzur M., 2008, ApJ, 685, 147

Norman C. et al., 2002, ApJ, 571, 218

Obreschkow D., Rawlings S., 2009, ApJ, 696, L129

Oliver S. J. et al., 2012, MNRAS, 424, 1614

Page K. L., O'Brien P. T., Reeves J. N., Turner M. J. L., 2004, MNRAS, 347, 316

Page M. J., Carrera F. J., 2000, MNRAS, 311, 433

Peacock J. A., 1983, MNRAS, 202, 615

Peterson B. M., 2003, An Introduction to Active Galactic Nuclei. Cambridge University Press

Pilbratt G. L. et al., 2010, A&A, 518, L1

Poglitsch A. et al., 2010, A&A, 518, L2

Popesso P. et al., 2009, A&A, 494, 443

Prescott M. K. M., Impey C. D., Cool R. J., Scoville N. Z., 2006, ApJ, 644, 100

Press W. H., Teukolsky S. A., Vetterling W. T., Flannery B. P., 1992, Numerical recipes in FORTRAN. The art of scientific computing

Pudritz R. E., Hardcastle M. J., Gabuzda D. C., 2012, Space Sci. Rev., 169, 27

Rafferty D. A., Brandt W. N., Alexander D. M., Xue Y. Q., Bauer F. E., Lehmer B. D., Luo B., Papovich C., 2011, ApJ, 742, 3

Ranalli P., Comastri A., Setti G., 2003, A&A, 399, 39

Ranalli P. et al., 2013, A&A, 555, A42

Ravikumar C. D. et al., 2007, A&A, 465, 1099

Rémy-Ruyer A. et al., 2013, A&A, 557, A95

Risaliti G., Elvis M., 2004, A panchromatic view of AGN

Rosario D. J. et al., 2013, ApJ, 771, 63

Rosario D. J. et al., 2012, A&A, 545, A45

Roseboom I. G. et al., 2012, MNRAS, 419, 2758

Roseboom I. G. et al., 2010, MNRAS, 409, 48

Rowan-Robinson M., Valtchanov I., Nandra K., 2009, MNRAS, 397, 1326

Rupke D. S. N., Veilleux S., 2011, ApJ, 729, L27

Salvato M. et al., 2009, ApJ, 690, 1250

Salvato M. et al., 2011, ApJ, 742, 61

Sanders D. B. et al., 2007, ApJS, 172, 86

Sandstrom K. M. et al., 2012, ApJ, 744, 20

Santini P. et al., 2009, A&A, 504, 751

Santini P. et al., 2012a, A&A, 538, A33

Santini P. et al., 2014, A&A, 562, A30

Santini P. et al., 2012b, A&A, 540, A109

Sazonov S. et al., 2012, ApJ, 757, 181

Schinnerer E. et al., 2007, ApJS, 172, 46

Schmidt M., 1959, ApJ, 129, 243

Schmidt M., 1968, ApJ, 151, 393

Schmidt M., Green R. F., 1983, ApJ, 269, 352

Schmidt M., Schneider D. P., Gunn J. E., 1995, AJ, 110, 68

Scoville N. et al., 2007, ApJS, 172, 1

Scoville N. et al., 2014, ApJ, 783, 84

Setti G., Woltjer L., 1989, A&A, 224, L21

Shankar F., Weinberg D. H., Miralda-Escudé J., 2009, ApJ, 690, 20

Shao L. et al., 2010, A&A, 518, L26

Shi Y., Helou G., Armus L., Stierwalt S., Dale D., 2013, ApJ, 764, 28

Silva L., Maiolino R., Granato G. L., 2004, MNRAS, 355, 973

Silverman J. D. et al., 2008, ApJ, 679, 118

Silverman J. D. et al., 2009, ApJ, 696, 396

Silverman J. D. et al., 2010, ApJS, 191, 124

Simpson C., 2005, MNRAS, 360, 565

Smith D. J. B. et al., 2012, MNRAS, 427, 703

Spergel D. N. et al., 2003, ApJS, 148, 175

Steffen A. T., Barger A. J., Cowie L. L., Mushotzky R. F., Yang Y., 2003, ApJ, 596, L23

Sturm E. et al., 2011, ApJ, 733, L16

Szokoly G. P. et al., 2004, ApJS, 155, 271

Taylor E. N. et al., 2009, ApJS, 183, 295

Tozzi P. et al., 2006, A&A, 451, 457

Treister E., Krolik J. H., Dullemond C., 2008, ApJ, 679, 140

Treister E., Urry C. M., 2006, ApJ, 652, L79

Treister E. et al., 2009, ApJ, 693, 1713

Tremonti C. A. et al., 2004, ApJ, 613, 898

Troncoso P. et al., 2014, A&A, 563, A58

Trump J. R. et al., 2009, ApJ, 696, 1195

Trump J. R. et al., 2007, ApJS, 172, 383

Turner T. J., George I. M., Nandra K., Mushotzky R. F., 1997, ApJS, 113, 23

Ueda Y., Akiyama M., Hasinger G., Miyaji T., Watson M. G., 2014, ApJ, 786, 104

Ueda Y., Akiyama M., Ohta K., Miyaji T., 2003, ApJ, 598, 886

Ueda Y. et al., 2008, ApJS, 179, 124

Urry C. M., Padovani P., 1995, PASP, 107, 803

van Dokkum P. G. et al., 2011, ApJ, 743, L15

Vanzella E. et al., 2008, A&A, 478, 83

Vanzella E. et al., 2010, ApJ, 725, 1011

Vignali C., Alexander D. M., Gilli R., Pozzi F., 2010, MNRAS, 404, 48

Vignali C., Mignoli M., Gilli R., Comastri A., Iwasawa K., Zamorani G., Mainieri V., Bongiorno A., 2014, ArXiv e-prints

Vito F., Gilli R., Vignali C., Comastri A., Brusa M., Cappelluti N., Iwasawa K., 2014a, ArXiv e-prints

Vito F. et al., 2014b, MNRAS, 441, 1059

Vito F. et al., 2013, MNRAS, 428, 354

Volonteri M., Begelman M. C., 2010, MNRAS, 409, 1022

Wang J.-M., Zhang E.-P., Luo B., 2005, ApJ, 627, L5

Whalen D. J., Fryer C. L., 2012, ApJ, 756, L19

Willott C. J. et al., 2009, AJ, 137, 3541

Willott C. J., McLure R. J., Jarvis M. J., 2003, ApJ, 587, L15

Wright N. J., Drake J. J., Civano F., 2010, ApJ, 725, 480

Wu X.-B. et al., 2010, A&A, 10, 737

Wuyts S., Labbé I., Schreiber N. M. F., Franx M., Rudnick G., Brammer G. B., van Dokkum P. G., 2008, ApJ, 682, 985

Wuyts S., van Dokkum P. G., Franx M., Förster Schreiber N. M., Illingworth G. D., Labbé I., Rudnick G., 2009, ApJ, 706, 885

Xue Y. Q. et al., 2011, ApJS, 195, 10

Yencho B., Barger A. J., Trouille L., Winter L. M., 2009, ApJ, 698, 380